THE UNIVERSITY
of EDINBURGH

Investigating the relationship between depression and cardiovascular diseases in three cohort studies

Regina Prigge

2020

The University of Edinburgh

A thesis submitted for the degree of

Doctor of Philosophy

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where explicitly stated otherwise the work presented is entirely my own.

Signature: ...……..…. Date: ……

Acknowledgements

First of all, I would like to thank my supervisors, Dr Caroline Jackson and Professor Sarah Wild, for their incredible support and guidance throughout my PhD. You have given up so much of your time to discuss my project, to provide feedback on earlier drafts of this thesis, to encourage me to follow my own ideas and intuition, and to introduce me to researchers from this university and abroad. I truly could not have asked for a better supervisor team to accompany me on this journey.

I am grateful to my work colleagues, many of whom have become friends along the way, for their advice and mental support. I have enjoyed working with all of you! I valued our exchange of ideas, I am thankful for feedback that I received during practice runs of presentations, and very much enjoyed our chats during lunchtime. I cannot wait to celebrate this PhD with all of you – I will bring the cake!

I would like to thank my parents, Birgit and Dieter, who provided me with the belief that I could achieve anything if I put my mind to it. Without you I would not be where I am today. Thank you for all your support and love.

Finally, I would like to thank my partner Jörn for cheering me on along the way. Your love, support and constant belief in me have helped so much! I cannot wait to start our next adventure.

Abstract

Background: Depression and cardiovascular diseases (CVD) are significant public health concerns. Whilst CVD are recognised as the leading cause of death worldwide, depression is responsible for a substantial non-fatal disease burden globally. A bidirectional association between depression and CVD has been reported in primary research and systematic reviews. However, there is an ongoing debate whether the evidence base is sufficient to acknowledge depression as an independent risk factor for subsequent CVD. Multiple potential mechanisms have been proposed but the exact nature of the association remains poorly understood.

Methods: I conducted a systematic review and meta-analysis to identify and critically appraise existing studies of the association between clinical depression or depressive symptoms and risk of major cardiovascular events (MCVE). To overcome some of the methodological shortcomings of existing studies I carried out quantitative analyses of three cohort studies. I performed a retrospective cohort study using Cox proportional hazard models to estimate the risk of MCVE among UK Biobank participants with depression, antidepressant use, hospital admission with depression, and self-reported depression relative to participants without the exposure of interest. Furthermore, I investigated the role of selected comorbidities and socioeconomic factors as potential effect-modifying factors. Using data from the Longitudinal Study on Women's Health (ALSWH), I used latent process mixed modelling to identify subgroups of participants with similar patterns of depressive symptoms over time and explored cardiovascular risk factor profiles at baseline and end of follow-up. Lastly, using data from the Whitehall II study, I compared psychological distress trajectories of individuals prior to diagnosis of cardiovascular events to psychological distress trajectories of individuals free from cardiovascular events over the same time period.

Results: The meta-analysis of 51 studies suggested that the risk of MCVE was higher among individuals with clinical depression or depressive symptoms relative to non-exposed individuals. However, results were influenced by methodological

shortcomings of existing studies. For example, the role of covariates as potential confounding or mediating factors was unclear and studies were prone to information bias since they relied on a single measure of clinical depression or depressive symptoms at baseline. Using data from the UK Biobank, different measures of depression were associated with increased risk of MCVE even after adjustment for a wide range of potential confounding factors. The increased risk was particularly high among individuals with comorbidities and among individuals from lower socioeconomic backgrounds. Using data from the ALSWH, three subgroups of women with distinct depressive symptom trajectories were identified. There were differences between groups with regards to profiles of key cardiovascular risk factors at baseline and end of follow-up. Women with stable moderate and fluctuating depressive symptoms had lower educational attainment, found it harder to manage on their income, reported more adverse health behaviours and medical conditions and gained weight more rapidly than women with stable low depressive symptoms. Using data from the Whitehall II study, there was some evidence of small differences between mean predicted psychological distress scores among those with and without cardiovascular events in the time prior to diagnosis, death, or end-of follow-up but no clear evidence for differences in the patterns of psychological distress over time.

Conclusions: This project addressed a number of shortcomings of the existing evidence base. The results of the UK Biobank analysis suggest that it is unlikely that residual confounding due to unmeasured covariates alone explained the observed associations between different measure of depression and MCVE. In addition, there was some evidence that the risk of MCVE might differ between subgroups of individuals with clinical depression or depressive symptoms. The analysis of ALSWH data highlighted the importance of considering repeat assessments of depressive symptoms. Furthermore, it illustrated that latent process mixed modelling might be a useful tool to identify potentially clinically meaningful groups with different patterns of depressive symptoms over time. The results of the Whitehall II analysis suggest that it is unlikely that reverse causation or overlap of symptoms

between subclinical CVD and somatic symptoms of depression explained the observed association.

Whilst a number of shortcomings of the existing evidence base could be addressed, some alternative explanations of the observed association remain to be investigated. For example, iatrogenic effects of psychotropic medications might partly explain the observed association between depression and CVD and there might be common risk factors including genetic predisposition for depression and CVD. Considering the substantial public health burden of both depression and CVD, it should remain a public health priority to further advance our understanding of the relationship between depression and subsequent CVD in future research.

Lay summary

People with depression have higher risks of cardiovascular diseases, such as stroke and heart attacks, than people without depression. However, we do not know why that is. It might be that depression itself alters the risk of stroke and heart attacks but there are a number of potential alternative explanations. For example, depression and stroke are both linked to low educational attainment which in turn might explain the link between depression and stroke.

This PhD aimed to improve our understanding of why individuals with depression might have higher risks of cardiovascular diseases using data from three large research studies. More specifically, I explored two potential alternative explanations of the link between depression and cardiovascular diseases. First, previous research did not consider all variables that are linked to both depression and stroke. This might have led to a spurious link between depression and cardiovascular disease in previous research. Second, instead of depression altering the risk of stroke, it might actually be the other way around. Even before someone is diagnosed with stroke or heart attack, there are undetected or undiagnosed changes in the body of that person. If these changes alter the risk of depression, it might explain the link between depression and cardiovascular diseases. A further shortcoming of previous research was that depression was often measured at one point in time. Since we know that depression comes and goes, I identified people with similar patterns of depressive symptoms over time. I then investigated if these groups differed with regard to social class, health behaviours such as smoking or physical inactivity, and medical conditions such as high blood pressure and diabetes.

In keeping with previous research, my analyses suggested that people with depression have higher risks of cardiovascular diseases than people without depression. Furthermore, people with depression who live in areas of high deprivation, who are less well educated, who have high blood pressure, diabetes, or high cholesterol levels were are at particularly high risk of stroke and heart attacks. In addition, I identified three groups with different patterns of depressive symptoms

over time. From the age of 20 to 40 years most people had stable low levels of depressive symptoms, a considerable number had stable moderate depressive symptoms, and some people had fluctuating depressive symptoms over time. These patterns of depression over time might be important. People with stable moderate and fluctuating depressive symptoms had lower educational levels and found it harder to manage on their income than people with stable low depressive symptoms. Furthermore, they reported more adverse health behaviours and had more medical conditions. The high risk of cardiovascular diseases among people with depression was not explained by variables that are linked to both depression and cardiovascular diseases. Also, undetected or undiagnosed changes in the body of people prior to diagnosis with stroke or heart attack did not seem to explain the altered risk of cardiovascular disease among people with depression.

This PhD addressed a number of shortcomings of previous research but some alternative explanations of the high risk of cardiovascular diseases among people with depression remain to be investigated. Since both depression and cardiovascular diseases affect a large number of people worldwide, it is important to conduct more research on this topic.

Table of Contents

List of Tables

List of Figures

Abbreviations

AIC	Akaike information criterion
ALSWH	Australian Longitudinal Study on Women's Health
BIC	Bayesian Information Criterion
BMI	Body mass index
CALIBER	Cardiovascular research using Linked Bespoke studies and Electronic Health Records
CASP	Critical Appraisal Skills Programme
CBVD	Cerebrovascular diseases
CES-D	Centre for Epidemiological Studies Depression scale
CHD	Coronary heart disease
CI	Confidence interval
CIDI-SF	Composite International Diagnostic Interview – Short Form
CPSS	Cohen Perceived Stress Scale
CRPD	Clinical Research Practice Datalink
CSE	Certificate of Secondary Education
CVD	Cardiovascular diseases
DALYs	Disability adjusted life years
DIS	Diagnostic Interview Schedule
DSM	Diagnostic and Statistical Manual of Mental Disorders
EPESE	Established populations for epidemiologic studies of the elderly
GDS	Geriatric Depression Scale
GEE	Generalised Estimating Equation
GHQ	General Health Questionnaire
GP	General Practitioner
HADS-D	Hospital Anxiety and Depression Scale – Depression subscale
HPL	Human Population Laboratory Depression Scale
HR	Hazard ratio
HRT	Hormone replacement therapy
ICD	International Classification of Diseases
IHD	Ischaemic heart disease
IMPACT	Improving Mood with Psychoanalytic and Cognitive Therapies
IQR	Interquartile range
IRR	Incidence rate ratio
ISD	Information Services Division
MAR	Missing at random
MCAR	Missing completely at random
MCVE	Major cardiovascular events
MDD	Major depressive disorder
MDE	Major depressive episode
MI	Myocardial infarction
MMPI OBD	Minnesota Multiphasic Personality Inventory – obvious depression subscale

MMPI-2	Second version of the Minnesota Multiphasic Personality Inventory
MMPI-2 D	Second version of the Minnesota Multiphasic Personality Inventory - Depression subscale in keeping with Hathaway & McKinley (1951)
MMPI-2 Dep	Second version of the Minnesota Multiphasic Personality Inventory - Depression subscale in keeping with Butcher et al (1989)
MNAR	Missing not at random
MONICA	Monitoring trends and determinants in cardiovascular disease project
MOPSY	MONICA psychosocial interview depression scale
NA	Missing value
NHS	National Health Services
NR	Not reported
NVQ	National Vocational Qualification
OR	Odds ratio
PHQ	Patient Health Questionnaire
PRIME-MD	Primary Care Evaluation of Mental Disorders
PY	Person years at risk
QOL	Quality of life
RCT	Randomised controlled trial
REGARDS	Reasons for Geographic and Racial Differences in Stroke study
RERI	Relative excess risk due to interaction
RR	Risk ratio
SAIL	Secure Anonymised Information Linkage
SCL-90	Symptom Checklist-90
SD	Standard deviation
SES	Socioeconomic status
SHR	Subdistribution hazard ratio
SSRI	Selective serotonin reuptake inhibitors
TCA	Tricyclic antidepressants
TIA	Transient ischaemic attack
UK	United Kingdom
US	United States of America
WHO	World Health Organisation
YLD	Years lived with disability
YLL	Years of life lost
Zung SDS	Zung Self-Rating Depression Scale

Chapter 1: Introduction

Depression and cardiovascular diseases (CVD) are significant public health concerns. Whilst CVD are recognised as the leading cause of death worldwide, depression is responsible for a substantial non-fatal disease burden globally (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016; GBD 2015 Mortality and Causes of Death Collaborators, 2016). Depressive disorders, which were assessed as a sub-category of all mental and substance use disorders in the Global Burden of Diseases Study, were attributable for the largest proportion of all mental and substance use disorder burden across all regions of the world (Whiteford et al, 2013).

A bidirectional association between depression and CVD has been observed in primary studies and systematic reviews. It is well established that physical diseases such as CVD increase the risk for depression. However, the relationship between depression and subsequent CVD remains controversial. Multiple potential mechanisms have been proposed, some of which propose causal pathways between depression and CVD, others offer alternative explanations (Penninx, 2016b). However, the exact nature of the association remains poorly understood (Fiedorowicz, 2014; Nemeroff & Goldschmidt-Clermont, 2012; Penninx, 2016a). If depression is indeed causally related to subsequent CVD, this will have important implications on resource allocation, and it offers possibilities for new preventive treatments for CVD if the pathophysiological effect of depression on arteries can be identified.

The aim of this project was to improve our understanding of the relationship between depression and subsequent CVD onset. In the following chapter this project is embedded in the wider context of existing literature (Chapter 2). The chapter gives an overview of depression and CVD, outlines their influence on the global burden of disease, and emphasises the importance of considering the comorbidity of mental and physical illnesses. Chapter 3 presents a systematic review and meta-analysis of the relationship between clinical depression or depressive symptoms and subsequent CVD. This review was conducted to identify and critically appraise existing research

and identify shortcomings of existing studies that could potentially be addressed as part of this project. Taking these shortcomings of existing studies into account, quantitative analyses of three cohort studies and linkage to administrative datasets were used to further our understanding of the relationship.

The first data analysis was based on the UK Biobank (Chapter 4). Researchers have highlighted that the observed association between depression and CVD might be due to residual confounding. Making use of unique advantages of the UK Biobank, in this project I explored the association between different measure of depression and major cardiovascular events (MCVE) whilst taking account of a wide range of potential confounding factors. Furthermore, I explored the role of selected comorbidities and socioeconomic factors as potential effect-modifying factors.

The second data analysis was based on the Australian Longitudinal Study on Women's Health (ALSWH) (Chapter 5). One major limitation of existing studies on the risk of CVD among individuals with depression is that a single measure of depression at baseline was used. Unlike many other cohort studies, the ALSWH has assessed depressive symptoms at multiple occasions over several years. Using group-based trajectory modelling, I aimed to identify subgroups of women with similar patterns of depressive symptoms over time. Furthermore, I investigated if these groups were characterised by different cardiovascular risk profiles at baseline and end of follow-up.

The third data analysis was based on the Whitehall II study (Chapter 6). One potential explanation for the association between clinical depression, depressive symptoms and risk of cardiovascular events is that symptoms of subclinical CVD overlap with somatic symptoms of depression. If the association between clinical depression, depressive symptoms and risk of cardiovascular events was indeed due to misclassification of individuals with subclinical CVD, one would expect an increase in depressive symptoms before diagnosis of CVD. Therefore, the aim of the work presented in this chapter was to compare depressive symptom trajectories of

individuals prior to diagnosis of cardiovascular events to depressive symptom trajectories of individuals free from cardiovascular events over the same time period.

In Chapter 7 the main findings of this thesis are presented, the contribution of each of the projects to our understanding of the relationship between depression and CVD are highlighted, strengths and limitations of this thesis are discussed, and implications for future research and practice are given.

The last chapter concludes this thesis (Chapter 8).

Chapter 2: Background

2.1 Depression

2.1.1 Introduction to depression

Major depression is a common mood (affective) disorder. Symptoms include low mood, loss of interest, sadness, poor sleep, and fatigue (American Psychiatric Association, 2013). Depending on the severity of the disorder, individuals may additionally have somatic symptoms, such as loss of libido or appetite, and symptoms may impair normal social and occupational functioning. Established risk factors of depression are stressful life events such as illness, divorce or bereavement, certain personality traits, loneliness, giving birth, and a family history of depression (National Health Service, 2016a). It is very likely that there is an interplay between different risk factors of depression which adds to the complexity of the disorder.

2.1.2 Assessment of depression

2.1.2.1 Clinical diagnosis

In routine health records most World Health Organisation (WHO) membership countries record a diagnosis of depression based on the International Classification of Diseases (ICD) (World Health Organisation, 2016a). The current official version is the ICD-10 Version: 2016 (World Health Organisation, 2016b). In this version, a differentiation is made between a single depressive episode and recurrent depressive disorder. Depending on the number and severity of symptoms, depressive episodes and recurrent depressive disorders can be rated as mild, moderate, or severe, and severe cases can present with or without psychotic symptoms. If depressive symptoms are chronic but not sufficiently severe or prolonged to diagnose a mild recurrent depressive disorder, patients may be diagnosed with dysthymia.

Another classification system for mental disorders is the Diagnostic and Statistical Manual of Mental disorders (DSM), published by the American Psychiatric Association (2013). The newest version of the DSM is DSM-5. In this version, depressive disorders are subdivided into disruptive mood dysregulation disorder, major depressive disorder (MDD), persistent depressive disorder (dysthymia),

premenstrual dysphoric disorder, substance/ medication induced depressive disorder, depressive disorder due to another medical condition, other specified depressive disorder, and unspecified depressive disorder. The category MDD most closely resembles the ICD diagnoses of depressive episode and recurrent depressive disorder. The DSM-5 describes nine diagnostic criteria that are used to determine the presence or absence of a MDD. The nine criteria are:

1. Depressed mood,
2. Diminished interest or pleasure,
3. Significant weight loss or weight gain, or decrease or increase in appetite,
4. Insomnia or hypersomnia,
5. Psychomotor agitation or retardation,
6. Fatigue or loss of energy,
7. Feeling of worthlessness or excessive or inappropriate guilt,
8. Diminished ability to think or concentrate, or indecisiveness, and
9. Recurrent thoughts of death (American Psychiatric Association, 2013).

In order to qualify for a depressive episode diagnosis, at least five out of nine criteria need to be present in the same two-week period and need to represent a change from previous functioning. In contrast to the ICD, a depressive disorder diagnosis further requires the presence of at least one out of two main criteria (depressed mood and/ or diminished interest or pleasure). If symptoms additionally cause significant impairment in important areas of the person's life, and the episode is not due to physiological effects of a substance or an underlying medical condition, the episode will be referred to as major depressive episode (MDE). In keeping with the ICD, a differentiation is made between mild, moderate and severe depressive episodes that can occur once (single episode) or multiple times (recurrent episodes). The DSM describes dysthymia as persistent depressive disorder with depressed mood being present for most of the day, for more days than not, for at least two years among adults. An individual may have both diagnoses at the same time, a MDD diagnosis alone, or dysthymia alone. Furthermore, the DSM allows to add specifiers for

depressive disorders, such as depressive disorder with melancholic or atypical features. Patients with a depressive disorder with atypical and melancholic features differ in their capacity of mood reactivity. Whilst patients with atypical depression might experience pleasure in the presence of positive events, patients with melancholic depression do not show improved mood in these situations (American Psychiatric Association, 2013). The DSM depression diagnosis is meant to be based on a semi-structured clinical interview administered by a clinician or trained mental health professional (First et al, 2015). The structured clinical interview for DSM-5 disorders was published as clinical version, research version, and clinical trials version.

Whilst the assessment of depression through a structured clinical interview based on DSM criteria is often seen as the gold standard for the assessment of depression, there also has been criticism of the criteria. The proposed diagnostic criteria of most mental disorders described in the DSM are based on diagnostic criteria proposed by a group of clinicians in 1972 (Feighner et al, 1972). The MDD criteria date back to a list of criteria published by Cassidy et al (1957). The group of clinicians working on the Feighner criteria changed the criteria proposed by Cassidy et al (1957) without giving a reason or empirical support (Zimmerman et al, 2015). As a result, some researchers have questioned the diagnostic reliability and validity of the proposed criteria, especially since only minor changes have been made in newer versions of the DSM (Fried & Nesse, 2015a; Lux & Kendler, 2010; Zimmerman et al, 2015). In addition, Zimmerman et al (2015) has highlighted that there are 227 ways to meet the nine DSM diagnostic criteria and two patients with the same diagnosis might not share a single symptom. Furthermore, Zimmerman et al (2015) emphasised that the large number of potential combinations led to immense diagnostic heterogeneity in practice. Of note, the researchers used a conservative approach by counting symptoms profiles according to the nine DSM criteria. Had they taken into account all sub-symptoms (e.g. weight loss and weight gain), there would have been 14,528 possible combinations. In a similar project, Fried & Nesse (2015a) investigated the degree of symptom heterogeneity by counting the number of unique DSM symptom profiles of

3,703 outpatients at the beginning of a randomised controlled trial (RCT). Similar to Zimmerman et al (2015), Fried & Nesse (2015a) found that only 2% of the sample had the most common symptom profile, and 14% of the participants showed unique symptoms profiles not shared with anyone else in the sample.

2.1.2.2 Rating scales

Rating scales measure the occurrence of a pre-specified set of depressive symptoms and/ or depressive symptom frequency. The severity of depressive symptoms is determined by adding up symptom scores to create a sum score. This sum score is usually compared against a pre-defined threshold that establishes the presence or absence of depressive symptoms. Rating scales are frequently used in epidemiological research because they require fewer financial resources, they can be used by lay interviewers, and they can assess the presence and/ or severity of depressive symptoms among individuals who do not use or have access to health care services.

A disadvantage of using rating scales is that different scales might measure different concepts of depressive symptomatology. It has been highlighted that different rating scales should not be used interchangeably (Skorikov & Vandervoort, 2003) since different rating scales collect information on different depressive symptoms, they were developed for the use in different samples, and they use different look-back periods. For example, the Centre for Epidemiological Studies Depression Scale (CES-D) was developed as an instrument to assess the frequency of 20 depressive symptoms in the general population in the past week (Radloff, 1977). In contrast, the Beck Depression Inventory was developed to measure the severity of 21 depressive symptoms in clinically depressed individuals at the present time (Beck et al, 1961). Other authors agree with the argument made by Skorikov & Vandervoort (2003). Fried (2017a) investigated the content overlap of seven common depression scales. Out of 125 items across seven scales the researcher identified 52 distinct depressive symptoms. Furthermore, 21 of the 52 depressive symptoms appeared in only one of the seven rating scales, and only six out of 52 symptoms were included in all seven

scales. Fried (2017a) concluded that the low content overlap might lead to research that is specific to the particular scale used. Additionally, he argued that the use of depressive symptom sum scores implies that all symptoms measure the same underlying concept. Considering how many different symptoms were identified, this assumption might not hold and depressive symptom sum scores might be a composite score of psychological problems rather than scores of depressive symptom severity (Fried et al, 2016). As a consequence, it has been argued that there should be more emphasis on individual depressive symptoms rather than on sum scores of rating scales (Fried & Nesse, 2015b; Fried et al, 2016). Whilst this is an interesting approach, it has remained challenging in practice since rating scales were neither developed nor validated to measure individual depressive symptoms (Fried et al, 2016).

Due to the range of different tools to assess depression and depressive symptoms, people who share related symptoms but have different aetiological and biological mechanisms might be grouped together (Navrady et al, 2018). This phenomenon is known as covert heterogeneity (Fried et al, 2014). Lux & Kendler (2010) investigated to what extent different depressive symptoms were associated with different clinical characteristics among patients with a DSM depression diagnosis. The researchers found that each of the nine unique DSM criteria showed varying associations with different comorbidities, demographic characteristics, and personality traits. The differences were partly explained by a distinction between cognitive and neurovegetative symptoms. However, it has proven difficult to identify subgroups that replicate in different populations. If there are indeed differences between subgroups, covert heterogeneity might have slowed down depression research. For example, Kapur et al (2012) argued that the lack of a biological test for depression might be due to the fact that one test cannot identify all subgroups of patients with depression. Similarly, Sullivan et al (2012) argued that covert heterogeneity hindered the identification of genetic variants as well as the replication of findings of genetic studies in different populations (see section 7.3.4.2 The need to investigate genetic pleiotropy in depression and CVD for a more detailed discussion).

2.1.3 The global burden of depression

Depression is responsible for a substantial burden globally. The Global Burden of Diseases Study provides estimates of disability adjusted life years (DALYs) which is the sum of all years of life lost (YLL) due to premature mortality and years lived with disability (YLD). In previous years, the global burden of depressive disorders was estimated as a sub-category of all mental and substance use disorders (Whiteford et al, 2013). Depressive disorders were further split into two distinct categories: MDD and dysthymia. In order to be captured within the study, cases had to meet the clinical definition according to the ICD or DSM criteria. Mental and substance use disorders were the fifth leading disorder category worldwide and the leading global cause of all non-fatal burden of disease in 2010 (Whiteford et al, 2013). Depressive disorders had the highest proportion of all mental and substance use disorder burden across all regions of the world. Whilst depressive disorders were the fourth leading cause of all YLD in 1990 and 2005, it has risen to the third leading cause of all YLD after lower back and neck pain and sense organ diseases in 2015 (GBD 2015 Disease and Injury Incidence and Prevalence Collaborators, 2016).

The latest estimates of the global burden of depressive disorders were published in 2017 (Table 2.1). In 2017 the global burden of depressive disorders was estimated as a subcategory of all mental disorders. Substance use disorders were treated as a separate category. 4.9% (95% CI: 3.9% – 5.9%) of all DALYs were due to mental disorders of which 26.5% and 8.2% were due to MDD and dysthymia, respectively (Global Burden of Disease Collaborative Network, 2018). All of the DALYs attributed to depressive disorders were due to YLD since neither MDD nor dysthymia were regarded as potential direct causes of deaths. The only mental disorders that were treated as potential direct causes of deaths were anorexia nervosa and bulimia nervosa. The burden of depressive disorders was higher among women than men (60.2% of all DALYs and 61.3% of all YLD) (Global Burden of Disease Collaborative Network, 2018).

*Table 2.1: Global burden of mental & depressive disorders in 2017 (Global Burden of Disease Collaborative Network, 2018)**

		Depressive disorders	
	Mental disorders (estimate and % of total)	**Major depressive disorder (estimate and % of total)**	**Dysthymia (estimate and % of total)**
DALYs	122.8 million (91.6 – 157.9) 4.9% (3.9% – 5.9%)	32.8 million (23.1 – 44.3) 1.3% (1.0% – 1.7%)	10.3 million (6.9 – 15.0) 0.4% (0.3% – 0.6%)
YLD	122.7 million (91.6 – 157.9) 14.4% (12.3% – 16.3%)	32.8 million (23.1 – 44.3) 3.9% (3.0% – 4.8%)	10.3 million (6.9 – 15.0) 1.2% (1.0% – 1.5%)
YLL	17.5 thousand (15.7 – 19.2) 0.0% (0.0% – 0.0%)	---	---

* Data are presented as n (95% confidence interval)

DALYs: Disability-adjusted life years, YLD: Years lived with disability, YLL: Years of life lost

Of note, these disability estimates capture short- or long-term health losses of the individual suffering from depression but do not capture losses due to effects on families or social and economic consequences (Whiteford et al, 2013). Furthermore, despite evidence of excess mortality attributable to multiple mental and substance use disorders (Liu et al, 2017), most of them were not recognised as underlying causes of death. Furthermore, suicides were captured under injuries (as self-harm). Since this might lead to an underestimation of the actual disease burden, Ferrari et al (2014) estimated suicide DALYs attributable to mental and substance use disorders. Mental and substance use disorders were responsible for 62.1% (95% CI: 43.8 – 75.3%) of all DALYs allocated to suicide in 2010. With 46.1% (95% CI: 28.0 – 60.8%) depression was responsible for the largest proportion. The inclusion of attributable suicide DALYs would have increased the YLL and, as a result, the global ranking of mental and substance use disorders would have changed from the fifth to third leading cause of burden in 2010. Even after consideration of the burden of suicide attributable to mental and substance use disorder the estimates might be underestimated "due to inadequate appreciation of the links between depression and other health conditions" (Charlson et al, 2011, p1), such as CVD.

2.2 Cardiovascular diseases

2.2.1 Introduction to cardiovascular diseases

CVD and cerebrovascular diseases (hereafter referred to as CVD) are two subgroups of circulatory diseases involving heart and blood vessels, as defined by the National Health Service (NHS) in the United Kingdom (UK) (National Health Service, 2018). Usually, the underlying disease of the blood vessel does not cause any symptoms. However, depending on the location of the vessels involved it can cause cardiovascular events such as myocardial infarctions (MI) or cerebrovascular events such as strokes (World Health Organisation, 2017). These events in turn cause symptoms such as pain or discomfort in the chest, arms, or shoulders, numbness of the face, difficulties walking, speaking, unconsciousness, or headaches. Major behavioural risk factors of CVD are smoking, physical inactivity, and an unhealthy diet, which in turn cause obesity, hypertension, and hyperlipidaemia. Both the behavioural risk factors as well as their consequences are recognised risk factors of CVD. CVD are largely preventable through primary prevention of the aforementioned behaviours or secondary prevention in those with established diseases (e.g. treatment with beta-blockers or statins) (National Health Service, 2018).

2.2.2 Assessment of cardiovascular diseases

In routine health records CVD diagnoses are often recorded based on the ICD. The current ICD-10 version differentiates between 100 diseases of the circulatory system (I00 to I99) which are grouped into the following ten categories (World Health Organisation, 2016b):

- Acute rheumatic fever
- Chronic rheumatic heart diseases
- Hypertensive diseases
- Ischaemic heart diseases
- Pulmonary heart disease and diseases of pulmonary circulation
- Other forms of heart disease
- Cerebrovascular diseases

- Diseases of arteries, arterioles and capillaries
- Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified
- Other and unspecified disorders of the circulatory system.

Due to the focus of this thesis, the assessment of cerebrovascular diseases (CBVD) and ischaemic heart diseases (IHD) will be described in more detail below.

CBVD are circulatory diseases involving blood vessels that supply the brain. CBVD are further divided into subarachnoid, intracerebral, and other non-traumatic intracranial haemorrhage, cerebral infarction, unspecified stroke, occlusion and stenosis of cerebral arteries not resulting in cerebral infarction, other CBVD such as cerebral aneurysm and sequelae of CBVD (World Health Organisation, 2016b). A differentiation between these disorders is usually made by a clinician based on the patient's symptoms and clinical signs, often accompanied by brain imaging findings. In 1970, the WHO defined stroke as "rapidly developing clinical signs of focal (at times global) disturbance of cerebral function, lasting more than 24h or leading to death with no apparent cause other than that of vascular origin" (Hatano, 1976, p541). It has been criticised that this definition has never been updated although major advances have been made in the understanding of aetiology, clinical recognition and diagnosis since the definition was formulated (Sacco et al, 2013). Although the term CBVD is meant to describe all of the aforementioned disorders, the terms stroke and CBVD are often used interchangeably. For example, although the American Heart Association called for an updated definition of stroke taking advances in science into account in 2013 (Sacco et al, 2013), they used the terms CBVD and stroke interchangeably in their publication of stroke statistics in 2019 (Benjamin et al, 2019). As a result, prevalence, incidence and mortality estimates may be less comparable between sources, and researchers need to pay particular attention to the definition used in existing literature.

IHD, also known as coronary heart diseases (CHD), are CVD involving blood vessels that supply the heart muscle. IHD are further classified into angina pectoris, acute

MI, subsequent MI, certain current complications following MI, other acute IHD, and chronic IHD. In 2012, the third universal definition of MI was published by members of the European Society of Cardiology, the American College of Cardiology Foundation, the American Heart Association, and the World Heart Federation (Thygesen et al, 2012). MI is defined as "myocardial cell death due to prolonged ischaemia" (Thygesen et al, 2012, p2023). Ischaemia is most commonly caused by atherosclerosis of arteries but in some cases it is caused by spasms of plaque free arteries or endothelial dysfunctions. The WHO adopted the same definition in high resource settings but incorporated more flexibility in diagnostic criteria in settings with resource constraints (Mendis et al, 2010).

2.2.3 The global burden of cardiovascular diseases

CVD are recognised as the leading causes of death worldwide (GBD 2015 Mortality and Causes of Death Collaborators, 2016). In 2017, CVD accounted for 31.8% (95% CI: 31.4% – 32.2%) of all deaths and 14.7% (95% CI: 13.6% – 15.6%) of all DALYs globally (Table 2.2). Although the age-standardised death rate has fallen over the past two decades, the absolute number of CVD deaths has risen due to ageing and growth of the world's population (Roth et al, 2015a). The majority of the CVD burden was attributable to both CBVD and IHD (Roth et al, 2015a).

The burden of CBVD was calculated as the sum of disease burdens due to ischaemic, haemorrhagic and other non-ischaemic strokes (Roth et al, 2015b). Stroke morbidity was calculated in accordance with the WHO definition of stroke, thus not counting transient ischaemic attacks (TIAs), epidural and subdural haemorrhage towards the stroke burden. In contrast, the burden due to stroke mortality was calculated after re-distributing all causes of death due to CBVD (ICD-10: I60 – 69) between ischaemic and haemorrhagic or other non-ischaemic strokes. Deaths due to TIA (ICD-10: G45) were coded as ischaemic strokes and deaths due to non-ruptured aneurysms (ICD-10: I67.0) were coded as haemorrhagic stroke. Nonspecific codes, including ICD-10 codes I64, I67.9, I68.8 and I69.4 – I69.9, were redistributed to ischaemic and haemorrhagic or other non-ischaemic strokes using a regression model (Roth et al,

2015b). Presumably, this was done under the assumption that causes of deaths were misclassified since disease classifications such as TIA or non-ruptured aneurysms should not have led to death. However, the paper describing the methodology for estimating the global burden of CBVD did not specifically state the rationale behind this approach (Roth et al, 2015b). CBVD were responsible for 36.1% of all CVD DALYs (Table 2.2). A large proportion of CBVD DALYs was due to YLL and a much smaller proportion was due to years lived with CBVD (85.8% and 14.2%, respectively). The CBVD burden was slightly larger in men than women (54.7% and 45.3% of all CBVD DALYs, respectively) (Global Burden of Disease Collaborative Network, 2018).

The burden of IHD was estimated as the sum of all YLL due to acute MI and IHD and all YLD with non-fatal acute MI, angina pectoris, and ischaemic heart failure (Moran et al, 2014). IHD were responsible for 46.5% of all CVD DALYs. As with CBVD, the contribution of YLL towards the overall burden of IHD was much higher than the contribution of YLD (96.9% and 3.1%, respectively). The burden of IHD was higher among men than women (61.7 and 38.3% of all IHD DALYs, respectively) (Global Burden of Disease Collaborative Network, 2018).

*Table 2.2: Global burden of cardiovascular diseases, ischaemic heart disease, and cerebrovascular diseases in 2017 (Global Burden of Disease Collaborative Network, 2018)**

	Cardiovascular diseases (reported estimate and % of total)	**Cerebrovascular Diseases (reported estimate and % of CVD)**	**Ischaemic Heart Disease (reported estimate and % of CVD)**
DALYs	365.9 million (355.2 – 376.7) 14.7% (13.6% – 15.6%)	132.1 million (126.5 – 137.4) 5.3% (4.9% – 5.7%)	170.3 million (167.1 – 174.0) 6.8% (6.3% – 7.4%)
YLD	35.7 million (26.4 – 45.5) 4.2% (3.7% – 4.6%)	18.7 million (13.6 – 23.7) 2.2% (1.7% – 2.6%)	5.3 million (3.7 – 7.2) 0.6% (0.5% – 0.7%)
YLL	330.2 million (324.9 – 335.2) 20.1% (19.7% – 20.4%)	113.4 million (111.0 – 116.2) 6.9% (6.7% – 7.1%)	165.0 million (162.3 – 168.6) 10.0% (9.8% – 10.2%)
Deaths	17.8 million (17.5 – 18.0) 31.8% (31.4% – 32.2%)	6.2 million (6.0 – 6.3) 11.0% (10.8% – 11.3%)	8.9 million (8.8 – 9.1) 16.0% (15.7% – 16.3%)

* Data are presented as n (95% CI)

CVD: cardiovascular diseases, DALYs: Disability-adjusted life years, YLD: Years lived with disability, YLL: Years of life lost

2.3 Mental-physical comorbidity

2.3.1 The relationship between major mental disorders and physical disorders

The co-existence of at least one mental and one physical health condition (hereafter referred to as mental-physical comorbidity) is a public health concern. In Scotland, prevalence of multimorbidity was estimated at 23.2% (95% CI: 23.1 – 23.2%) (Barnett et al, 2012), and mental-physical comorbidity accounted for 40% of all multimorbidity among primary care patients (McLean et al, 2014). McLean et al (2014) further highlighted that the prevalence of mental-physical comorbidity in the most deprived decile of the Scottish population was almost twice that of the least deprived decile (5.9%, 95% CI: 5.8 – 6.0%, and 11.0%, 95% CI: 10.9 – 11.2%, respectively). Furthermore, the onset of multimorbidity was ten years earlier among people living in the most deprived area than among people living in the least deprived area (Barnett et al, 2012). In the United States (US), the prevalence of mental-physical comorbidity was estimated at 9.5% among women and 6.0% among men (Bobo et al, 2016). Whilst this appears to be much lower than in Scotland, a comparison of estimates across populations is complicated due to the use of different data sources and the inclusion of different disorders in their analyses. Whilst Bobo et al (2016) took 15 conditions, including five mental health conditions into account, Barnett et al (2012) and McLean et al (2014) used information on 40 conditions including eight mental health conditions.

Individuals with major mental disorders die prematurely. Reisinger Walker et al (2015) performed a systematic review on the risk of all-cause mortality among people with a diagnosed mental disorder relative to people with no diagnosed mental disorder. Out of 148 included cohort studies 135 studies showed increased risks of all-cause mortality. The pooled risk ratio (RR) was estimated at 2.22 (95% CI: 2.12 – 2.33) and individuals with mental disorders had a median loss of ten years of potential life. Furthermore, there was some indication that the gaps between mortality rates and life expectancies might be widening (Lawrence et al, 2013; Saha et

al, 2007). An exception thereof is the study by Wahlbeck et al (2011) that shows a modest decrease in the life expectancy gap in Finland and Denmark and among women in Sweden. One potential reason for the excess mortality is lack of effective interventions that target risk factors of excess mortality in people with severe mental disorder such as behavioural risk factors, or screening and management of physical health conditions (Liu et al, 2017). Other potential reasons for the excess mortality are differences in treatment and/ or under-treatment of physical disorders in people with mental disorders (Newcomer & Hennekens, 2007; Reisinger Walker et al, 2015), and higher prevalence of adverse lifestyle factors such as smoking, physical inactivity, high caloric intake or poor diet quality (Liu et al, 2017; Newcomer & Hennekens, 2007; Reisinger Walker et al, 2015; World Health Organisation, 2015). Whilst potential reasons for the excess mortality have been proposed, the relative importance of different potential reasons remains to be established. Since the reasons are likely to be multifactorial, a multilevel model of interventions has been proposed to reduce excess mortality among people with severe mental disorders (Liu et al, 2017). This model combines individual-focused interventions, health-system focussed interventions, and community level and policy-focused interventions. Individual-focused interventions include but are not limited to smoking cessation programmes, behavioural weight management programmes aiming to improve diet and increase physical activity, and interventions addressing alcohol and substance abuse. Health-system focussed interventions address factors related to service delivery. Community level and policy-focused interventions include programmes addressing social support, stigma, employment, housing, and social welfare. Importantly, while suicides and accidents accounted for some of the excess mortality risk, 67.3% of all deaths among mental disorder patients were due to natural causes (Reisinger Walker et al, 2015). Furthermore, it has been emphasised that the main causes of deaths were similar to the general population, with CVD being the leading cause of death (Crump et al, 2013; Gatov et al, 2017; Nordentoft et al, 2013).

There is evidence that major mental disorders, defined as schizophrenia, bipolar disorder, and/ or MDD, are associated with CVD. In a large-scale meta-analysis,

Correll et al (2017) estimated the incidence and mortality from CVD among people with severe mental disorders. Individuals with severe mental disorders were at increased hazard of CVD incidence (pooled HR: 1.78, 95% CI: 1.60 – 1.98), cardiovascular related death (pooled HR: 1.85, 95% CI: 1.53 – 2.24), CHD (pooled HR: 1.54, 95% CI: 1.30 – 1.82), CBVD (pooled HR: 1.64, 95% CI: 1.26 – 2.14), and congestive heart failure (pooled HR: 2.10, 95% CI: 1.64 – 2.70), relative to individuals without any major mental disorder. Due to the focus of this thesis, the evidence on the association between depression and CVD will be discussed in detail in section 2.3.2 The relationship between depression and cardiovascular diseases. For both bipolar disorder and schizophrenia a scoping review was performed to identify existing meta-analyses on their association with subsequent CVD. In this review, three meta-analyses on the association between bipolar disorder and subsequent CVD and four meta-analyses on the association between schizophrenia and subsequent CVD were identified (Table 2.3). Both bipolar disorder and schizophrenia were associated with increased risks of CVD, CVD-related mortality, CBVD, stroke, and coronary heart failure (Correll et al, 2017; Fan et al, 2013; Li et al, 2014; Prieto et al, 2014). Whilst bipolar disorder was associated with increased risk of hypertension but not CHD (Ayerbe et al, 2018; Correll et al, 2017), schizophrenia was associated with increased risk of CHD but not hypertension (Ayerbe et al, 2018; Correll et al, 2017). Furthermore, individuals with bipolar disorder were at increased risk of MI but the confidence interval (CI) overlapped with the null value (Prieto et al, 2014). No meta-analysis on the association between schizophrenia and MI was identified. Meta-analyses on the association between antidepressant and antipsychotic drug use and risk of CVD are discussed in section 7.3.4.1 The need to disentangle the role of depression and psychotropic medications.

*Table 2.3: Existing meta-analyses on the association between bipolar disorder or schizophrenia and subsequent cardiovascular diseases (reverse chronological order)**

	Author	Outcome	Metric	Pooled estimate (95% CI)
Bipolar disorder	Ayerbe et al (2018)	Hypertension	IRR	1.27 (1.15 – 1.40)
	Correll et al (2017)	Cardiovascular diseases	HR	1.57 (1.28 – 1.93)
		Coronary heart disease		1.16 (0.76 – 1.78)
		Cerebrovascular diseases		1.60 (0.99 – 2.57)
		Coronary heart failure*		2.27 (1.49 – 3.45)
		Cardiovascular disease-related death		1.65 (1.10 – 2.47)
	Prieto et al (2014)	Myocardial infarction	RR	1.09 (0.96 – 1.24)
		Stroke		1.74 (1.29 – 2.35)
Schizophrenia	Ayerbe et al (2018)	Hypertension	IRR	0.94 (0.75 – 1.14)
	Correll et al (2017)	Cardiovascular diseases	HR	1.95 (1.41 – 2.70)
		Coronary heart disease		1.59 (1.08 – 2.35)
		Cerebrovascular diseases		1.57 (1.09 – 2.25)
		Cardiovascular disease-related death		2.45 (1.64 – 3.65)
	Li et al (2014)	Stroke (fatal or non-fatal)	RR	1.50 (1.25 – 1.80)
		Stroke (fatal)		1.65 (1.31 – 2.08)
	Fan et al (2013)	Cardiovascular diseases	RR	1.53 (1.27 – 1.86)
		Coronary heart disease		1.53 (1.27 – 1.86)
		Stroke		1.71 (1.19 – 2.46)
		Coronary heart failure		1.81 (1.42 – 2.29)

* Results are based on one study

CI: Confidence interval, HR: Hazard ratio, IRR: Incidence rate ratio, RR: Risk ratio

2.3.2 The relationship between depression and cardiovascular diseases

2.3.2.1 The comorbidity of depression and CVD

Depression is more common among people suffering from CVD than among the general population. The global prevalence of depressive disorders was estimated at 3.6% (95% CI: 3.3 – 3.9) in the general population in 2017, with a higher prevalence among women than men (4.4%, 95% CI: 4.1 – 4.7 and 2.8%, 95% CI: 2.6 – 3.1, respectively) (Global Burden of Disease Collaborative Network, 2018). Air et al (2016) investigated data on the prevalence of depression among cardiac disease patients and found that depression disproportionally affected people suffering from CVD with little variation between CVD subtypes. The authors concluded that at least one in five patients with CVD met the criteria for major depression (Air et al, 2016). Combining data from four diagnostic interview studies, Doyle et al (2015) reported that 20% of women and 12% of men with MI had major depression. Using interview and

questionnaire studies, this estimate increased to 36% of women and 29% of men reporting elevated depressive symptoms or major depression. Bobo et al (2016) investigated the coexistence of chronic mental and somatic disorders. Due to its high prevalence depression was the condition most commonly involved in mental-physical comorbidity. Hyperlipidaemia, hypertension, diabetes, cardiac arrhythmia, and arthritis were the most common somatic conditions co-occurring with depression.

2.3.2.2 Longitudinal studies

2.3.2.2.1 Depression among individuals with existing cardiovascular disease

Depression is associated with increased risks of mortality and recurrent non-fatal events among individuals with existing CVD. In a systematic review and individual patient data meta-analysis, Doyle et al (2015) reported an increased hazard of all-cause mortality among depressed MI patients relative to non-depressed MI patients. The association was stronger in men than women (pooled HR: 1.38, 95% CI: 1.30 – 1.47, and pooled HR: 1.22, 95% CI: 1.14 – 1.31, respectively), which partly might have been explained by disease severity. Nicholson et al (2006) performed a meta-analysis assessing depression as a prognostic factor among CHD patients. Twenty-four of the included primary studies defined a worse prognosis by assessing the risk of all-cause mortality among depressed CHD patients relative to non-depressed CHD patients. The pooled relative risk of dying from any cause was 1.80 (95% CI: 1.46 – 2.22) times higher in those with depression than in those without depression. Furthermore, the pooled relative risk of cardiovascular or cardiac mortality among depressed participants with CHD relative to non-depressed participants was estimated at 2.29 (95% CI: 1.33 – 3.94) using information from six primary studies. In addition, depression has been shown to be associated with a negative effect on functional outcomes. For example, poststroke depression is associated with worse quality of life (QOL), poor life satisfaction, and less efficient use of rehabilitation services (Kutlubaev & Hackett, 2014; Towfighi et al, 2017). Although the results of the individual studies included in these reviews were heterogeneous in terms of

participants included, measurement of depression, and covariates included, an expert group recommended "that the American Health Association should elevate depression to the status of a risk factor for adverse medical outcomes in patients with acute coronary syndrome" in 2014 (Lichtman et al, 2014, p1350). This was based on strong, consistent associations between depression and acute coronary syndrome and plausible biological mechanisms. However, it was emphasised that this recommendation was given despite a lack of evidence that treatment of depression lowers the risk of acute coronary syndrome.

2.3.2.2.2 Depression among individuals with no history of cardiovascular disease

Depression is associated with increased risk of CVD among individuals with no history of CVD. Thirteen meta-analyses on the association between clinical depression and/ or depressive symptoms and subsequent CVD were identified in a scoping review (Figure 2.1). These meta-analyses provided 37 pooled effect estimates on the association between different measures of depression and subsequent CVD, IHD, CBVD, stroke, ischaemic stroke, haemorrhagic stroke, MI, congestive heart disease, CVD-related mortality, and hypertension. Three meta-analyses reported estimates assessing depression based on depressive symptom rating scales (Charlson et al, 2013; Gan et al, 2014; Rugulies, 2002), five meta-analyses reported estimates using a measure of clinical depression (Barlinn et al, 2014; Charlson et al, 2013; Correll et al, 2017; Gan et al, 2014; Rugulies, 2002), and twelve meta-analyses reported estimates combining information obtained through rating scales and measures of clinical depression (Barlinn et al, 2014; Charlson et al, 2013; Dong et al, 2012; Gan et al, 2014; Leung et al, 2012; Li et al, 2015; Meng et al, 2012; Nicholson et al, 2006; Pan et al, 2011b; Rugulies, 2002; Wu & Kling, 2016; Wulsin & Singal, 2003). All meta-analyses indicated that depression is associated with increased risk of CVD but two CIs of smaller meta-analyses overlapped with the null value (Charlson et al, 2013; Gan et al, 2014). The pooled effect estimates were broadly consistent across studies with the exception of two studies on the risk of IHD among individuals with clinical depression relative to those without clinical depression that showed very high effect

estimates (Rugulies, 2002; Wu & Kling, 2016). Among the three studies that provided estimates on both depressive symptoms alone and clinical depression alone two studies provided higher effect estimates among individuals with clinical depression (Charlson et al, 2013; Rugulies, 2002), whereas the effect estimate among individuals with depressive symptoms was higher in the study by Gan et al (2014).

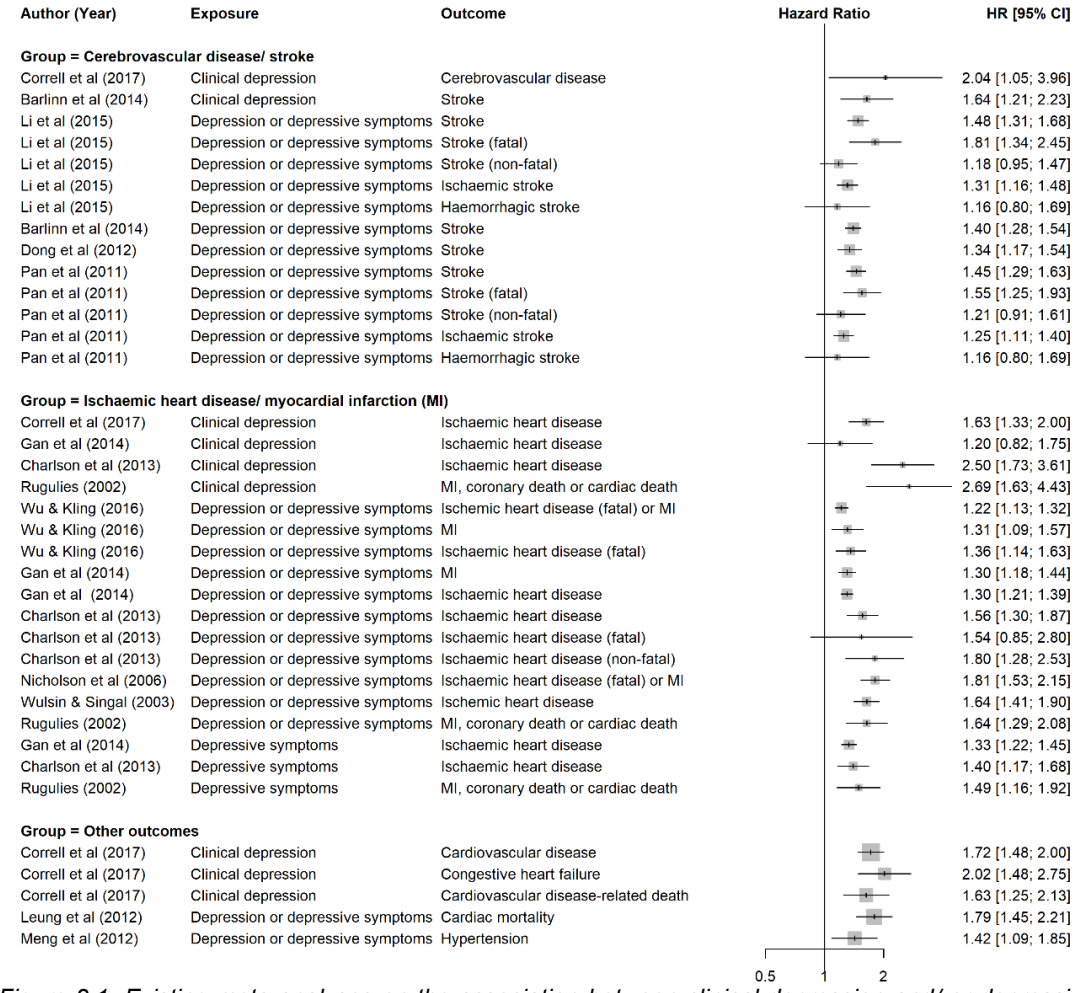


Figure 2.1: Existing meta-analyses on the association between clinical depression and/ or depressive symptoms and subsequent cardiovascular diseases

2.3.2.2.3 Limitations of existing observational studies

These associations suggest a close relationship between depression and CVD but they do not necessarily indicate a causal relationship. There is an ongoing debate about whether or not there is sufficient evidence to classify depression as independent risk factor for CVD. Whilst some argue that there is sufficient evidence to consider depression as independent risk factor for CVD (Fiedorowicz, 2014) and CHD

(Charlson et al, 2011; Goldston & Baillie, 2008), others argue that depression is yet to be confirmed as independent risk factor for CVD (Baxter et al, 2011) and CHD (Nicholson et al, 2006). Those that argue that depression is not yet established as independent risk factor for CVD point towards shortcomings of existing reviews and primary studies. First, existing reviews have struggled to explain the high heterogeneity between studies and methodological shortcomings of included studies have likely influenced the conclusions of existing reviews. For example, there is no agreement on how depression and CVD should be assessed. As a result, there are differences in exposure and outcome assessments across primary studies, which partly explains the high heterogeneity. In addition, most primary studies included in the systematic reviews have assessed depression once at baseline. Since depressive disorders are episodic and fluctuating, assessing the presence of clinical depression or severity of depressive symptoms at one point in time might not capture the longitudinal trajectory of the disorder (Colman & Ataullahjan, 2010). Additionally, longitudinal observational studies are needed that provide valuable insights into sequence of events and changes over time among participants enrolled in the study. Third, many primary studies inadequately adjusted for potential confounding or effect-modifying factors (Nicholson et al, 2006). Lastly, publication bias has been detected (Wu & Kling, 2016) with an indication that studies with negative results have not been published. In contrast, other researchers, who argue in favour of sufficient evidence to assess depression as an independent risk factor of CVD, emphasise that a number of the Bradford-Hill criteria for causality are met (Hill, 1965). First, the association between depression and CVD has been consistent and strong in existing primary studies (Charlson et al, 2011; Goldston & Baillie, 2008). Moreover, it was stated that there is sufficient evidence for a temporal relationship between depression and CVD since multiple well-conducted prospective studies have excluded people with a known history of CVD at baseline, and evidence for a dose-response relationship between depression and CVD was observed (Charlson et al, 2011; Goldston & Baillie, 2008). In contrast to the opposing view, these researchers argued that most studies sufficiently controlled for confounding, particularly because it

could be argued that some covariates might be on the causal pathway. Additionally, an association has been observed across multiple somatic and mental health conditions suggesting that there is biological plausibility for a causal relationship. Lastly, it was argued that possible biological and behavioural pathways have been identified in the relationship between depression and CVD.

2.3.2.3 Proposed mechanisms

Several pathways might explain the increased risk of CVD among people suffering from depression. Potential mediating factors, such as biological dysregulations, behavioural factors, and worse medical care causally relate depression to CVD, whereas the residual confounding hypothesis, the vascular depression hypothesis, the common soil hypothesis, and iatrogenic effects of medication could contribute to a non-causal relationship between depression and CVD. It should be noted that all of these explanations potentially play a role, which adds to the complexity of the topic.

The main candidate pathways for biological dysregulations are dysregulations of the autonomic nervous system, hypothalamus-pituitary-axis hyperactivity, inflammatory processes, endothelial dysfunctions, and alterations of the platelet-clotting cascade. The autonomic nervous system has two divisions, one of which regulates the fight or flight response (sympathetic nervous systems) whereas the other maintains bodily functions such as digesting and resting (parasympathetic nervous system). In stressful situations the sympathetic nervous system is more active than the parasympathetic nervous system. Due to chronic stress, it has been hypothesised that the sympathetic nervous system is over-active among depressed patients. For example, it has been observed that heart-rate variability, an indirect measure of autonomic nervous system functioning, is decreased among depressed patients (Kemp et al, 2010). Since autonomic dysregulations are associated with increased CVD risk (Licht et al, 2013), it was proposed as a possible mediator of the relationship between depression and CVD. Since the hypothalamus-pituitary-adrenal-axis (HPA axis), a neuroendocrine system consisting of the hypothalamus, pituitary and adrenal glands, is involved in the stress response of the body, it has

been proposed as a second potential mediator of the relationship between depression and CVD (Fiedorowicz, 2014). A hyperactive HPA axis leads to increased cortisol and catecholamine levels, both of which have been observed in depressed patients (Stetler & Miller, 2011). Highly increased cortisol levels influence platelet activation and aggregation, and alterations in these processes have been observed among depressed patients (Nemeroff & Goldschmidt-Clermont, 2012). Since platelet activation and aggregation are involved in arterial repair, it has been hypothesised that the alteration of these processes might be responsible for the failure to repair arteries which has been observed among people with depression (Nemeroff & Goldschmidt-Clermont, 2012). Additionally, inflammatory processes, which limit the function of the arterial wall, have been observed more often among people with depression than in the general population (Beatriz Currier & B Nemeroff, 2010). Through the combination of impaired arterial function and failure to repair arteries, depression might lead to arteriosclerotic processes, which in turn lead to CVD. However, Nemeroff & Goldschmidt-Clermont (2012) and Fiedorowicz (2014) criticise that the mechanisms are mostly speculative and the precise nature of this relationship is yet to be established. Whooley & Wong (2013) argue that the previously described biological changes may result from poor health behaviours, which have been observed among depressed patients.

Behavioural factors that might act as mediators of the relationship between depression and CVD include physical inactivity, CVD medication non-adherence, excessive alcohol use, and smoking. Win et al (2011) investigated the influence of adding physical activity variables to a Cox proportional hazards model investigating the influence of depressive symptoms on cardiovascular mortality. They observed a reduction of the log hazard ratio (HR) by 26% and concluded that physical activity acts as a partial mediator of the depression–CVD mortality association. Another study investigated the potential synergistic association between depression and smoking with heart disease and found the strongest association among depressed adults with greater smoking exposure compared with less smoking exposure (Carroll et al, 2017). Whooley et al (2008) aimed to determine why there is an association

between depression and cardiovascular events among patients with CHD and investigated the change in the strength of the age-adjusted log HRs after adjustment for potential confounding and mediating factors. They concluded that the association between depression and cardiovascular events was largely explained by behavioural factors since adjustment for alcohol use, medication non-adherence, smoking, and physical activity resulted in 5%, 10.9%, 5.3%, and 31.7% reduction of the age-adjusted log HR, respectively. However, it was not described clearly whether these adjustments were done individually or incrementally.

Depression might be associated but not causally related to CVD due to residual confounding in existing research, shared risk factors, reverse causation, or iatrogenic effects of antidepressants. The residual confounding hypothesis originated from findings that depressed patients are markedly different from non-depressed patients in terms of their sociodemographic characteristics as well as lifestyle and biological factors (Penninx, 2016a). Consequently, these factors should ideally be taken into account in statistical analyses. Although most primary studies have presented adjusted results, some studies have not adjusted their results for main confounding factors. Due to the strength and consistency of the association some argue that it is unlikely due to residual confounding alone (Penninx, 2016a). Others emphasise the lack of adjustments in primary research (Nicholson et al, 2006) and argue that variation of estimates might be due to variation in statistical adjustment (Stapelberg et al, 2013). A second hypothesis is that depression and CVD share common roots (common soil hypothesis), but the effect of these common roots manifests earlier on one disease than the other, thus leading to a spurious longitudinal association between depression and CVD. Examples of potential shared risk factors are shared genetic variants (genetic pleiotropy) (de Geus, 2006), early-life factors or low socioeconomic status (SES) (Mortimer et al, 2016). Furthermore, underlying cerebrovascular disease itself might "predispose, precipitate, or perpetuate some geriatric depressive symptoms" (Alexopoulos et al, 1997, p915). This hypothesis is known as vascular depression hypothesis. It originated from findings that both

individuals with geriatric depressive symptoms and individuals with vascular dysregulations show greater severity of white matter lesions (ischaemic lesions in the brain), identified as white matter hyperintensities on MRIs (Taylor et al, 2013). If depressive symptoms are indeed related to worsening of underlying vascular disease, depression may not be causally related with subsequent cardiovascular events but instead reflect a prodromal feature of worsening vascular disease itself. A third hypothesis is that psychotropic medications and not depression itself are responsible for the increased CVD risk among people with depression. However, it has been challenging to disentangle the effects of depressive disorders and antidepressants since antidepressant use is likely to act as an identifier of disease severity at the same time (see section 7.3.4.1 The need to disentangle the role of depression and psychotropic medications for a more detailed discussion). Also, in the literature on the association between antidepressants and CVD risk methodological shortcomings similar to those observed in the literature between depression and CVD have been reported (Shin et al, 2014).

It remains to be established to what extent each proposed mechanism contributes to the increased risk of CVD among people with depression. Due to the complexity of the relationship between depression and CVD, the likely interplay of different mechanisms, and challenges in using causal inference methods in research on psychological disorders (see section 7.3.3.3 Causal inference in research on psychological disorders for further discussion), quantifying the mediating effect of one or all of the mechanisms has proven challenging. Therefore, whilst potential mechanisms have been proposed, the importance of different mechanisms has yet to be established.

2.3.2.4 Intervention studies

2.3.2.4.1 Existing randomised controlled trials

Existing meta-analyses aimed at investigating the effectiveness of psychological interventions on risk of all-cause mortality, CVD-specific mortality, and recurrent non-fatal events among patient with a history of CVD have been inconclusive (Albus

et al, 2019; Linden et al, 2007; Reavell et al, 2018; Richards et al, 2017; Welton et al, 2009; Whalley et al, 2014). The most recently published systematic review and meta-analysis investigated the additional impact of psychological lifestyle change or distress management interventions on depression, anxiety, QOL, cardiac morbidity, and cardiovascular or total mortality compared to usual care among patients with coronary artery disease or congestive heart failure (Albus et al, 2019). Whilst pooled effect estimates indicated a trend towards decreased depressive symptoms and reduced cardiac morbidity among those allocated to the intervention group, there was no evidence for changes in anxiety, QOL, and cardiovascular or all-cause mortality. These results are only partially consistent with another meta-analysis aimed at assessing the effectiveness of psychological interventions compared to usual care among CHD patients (Richards et al, 2017). Whilst neither the risk of all-cause mortality nor the risk of non-fatal MI was significantly different between the intervention and control group, there was a 21% (95% CI: 2% – 37%) reduced risk of cardiac mortality among the intervention group. Importantly, the authors of the meta-analyses highlighted the low quality of evidence and emphasised significant heterogeneity between studies due to differences in study populations and characteristics of interventions. Furthermore, it was highlighted that only 12 of 35 included RCTs used the existence of psychopathology at baseline as eligibility criterion. Goldston & Baillie (2008) emphasised that this has likely led to floor effects since participants had little need for psychological intervention thereby not benefitting from it. In addition, it was highlighted that a further potential explanation for the null findings might be that participants in these trials had already developed CVD (Stewart et al, 2014).

There is a lack of trials on the effect of depression treatment on first-ever CVD events. I have only identified one RCT that investigated the effectiveness of a psychological intervention among patients free from CVD at baseline. Stewart et al (2014) conducted an analysis using data from the Improving Mood- Promoting Access to Collaborative Treatment (IMPACT) RCT, in which they assessed the effect of collaborative care

compared to usual care among participants with major depression or dysthymia with and without established CVD at baseline. They observed a significantly decreased hazard of CVD, defined as fatal or non-fatal MI or stroke, among participants without baseline CVD (HR: 0.52, 95% CI: 0.31 – 0.86) but not among participants with baseline CVD (HR: 1.19, 95% CI: 0.70 – 2.03). However, the authors acknowledged limitations such as a post-hoc nature, low statistical power, and a non-effective psychological intervention among participants with baseline CVD. As a result, they concluded that there is a need for well-powered RCTs to replicate their findings and to elucidate mechanisms underlying the effect of depression on CVD.

2.3.2.4.2 Limitations of existing randomised controlled trials

There are major shortcomings in existing RCTs. First, the findings of existing studies might have been influenced by chance. The sample sizes of studies included in the systematic review by Albus et al (2019) ranged from 70 to 1,127 participants. Likely explanations of small sample sizes were high costs involved in running a trial and resource intensive treatments in trials involving psychological interventions. Furthermore, it has been highlighted that power calculations might have been based on existing studies that were too small to observe an effect or existing studies that were not comparable methodologically (Pedersen & Doyle, 2019). Due to small sample sizes true differences between groups might have remained undetected.

Second, selection bias is likely to have influenced the results of existing RCTs on the effect of psychological interventions on risk of cardiovascular events. Participants of RCTs underlay strict inclusion and exclusion criteria in order to increase homogeneity among trial participants. As a result, the limited number of potentially eligible participants were highly selected which affected the generalisability of findings. Furthermore, selection bias is likely to have influenced results due to self-selected samples. Participants of trials are usually highly motivated individuals who tend to have less serious medical conditions. In keeping with that it has been shown that individuals with mental health conditions are less likely to enrol in RCTs (Pogosova et al, 2015). As described above, one potential consequence of that was that enrolled

participants might have had little need for psychological interventions thereby not benefitting from it (Goldston & Baillie, 2008).

A further limitation of existing RCTs is that psychological interventions were not well-defined. A major shortcoming was that usual care tended to include advice on lifestyle and psychological advice. Whilst it might be unethical not to give this advice to the control group, the levels of treatment might not be sufficiently different between the intervention and control groups. Also, if usual care is effective, an additional psychological intervention might not have an additional effect (Albus et al, 2019; Pedersen & Doyle, 2019). Moreover, the use of different versions of treatments and comparators across studies complicated the comparison of findings of different RCTs. Observed differences across trials might have been due to varying effects of different psychological interventions or due to methodological differences across studies. To improve the quality and completeness of descriptions of interventions and enhance interpretation of observed differences across trials, it is recommended that future studies follow the template for intervention description and replication (TIDieR) checklist and guide (Hoffmann et al, 2014).

2.4 Summary of justification for this project

Depression and CVD both substantially contribute to the global burden of diseases. An association between depression and CVD onset and progression has been observed in numerous primary studies and existing reviews but there is an ongoing debate whether or not the evidence base is sufficient to acknowledge depression as an independent risk factor for CVD. Multiple plausible mechanisms have been proposed but the exact nature of the association between depression and CVD remains poorly understood. If depression is indeed a causal risk factor for CVD, this will have important implications on resource allocation, depression and CVD treatment, and it offers possibilities for new preventive treatments for CVD if the pathophysiological effect of depression on arteries can be identified. Therefore, the overall aim of this project is to further our understanding of the relationship between depression and subsequent CVD onset.

Chapter 3: The association between clinical depression or depressive symptoms and major cardiovascular events – A systematic review and meta-analysis

3.1 Background

Existing systematic reviews and meta-analyses have provided pooled effect estimates of the association between depression and subsequent CVD. However, they did not provide sufficient information to gain a detailed understanding of the literature and its methodological shortcomings. Whilst pooled effect estimates will be provided in this chapter to synthesise the results of existing studies, this review's primary aim was to provide a detailed assessment of the current evidence base. A particular focus of this review was to assess the strengths and weaknesses of existing studies and to identify whether studies have previously explored mechanisms and mediating pathways of the depression-CVD association. After completion of the review, I investigated the extent to which shortcomings of existing studies can be overcome as part of this project. Studies that have previously explored mechanisms and mediating pathways were identified because at the outset of this PhD one of the objectives was to explore mechanisms and mediating pathways between depression and CVD. However, due to methodological considerations and time restrictions this then was not taken further. A more detailed discussion of this methodological consideration is given in section 7.3.2 The need for an accurate definition of depression.

A preliminary search highlighted that a large number of primary studies on the association between depression and different CVD subtypes have been published. Given the vast amount of literature in this area a systematic review of all primary studies on the association between depression and any CVD subtypes was not deemed feasible within the time restrictions of this project. Therefore, it was decided to restrict this review to studies on the risk of first-ever fatal or non-fatal stroke or MI. Stroke and MI were chosen as outcomes because they are two common and severely disabling CVD events. Also, stroke and MI events are likely to be registered in

hospital records which was an important consideration because it had already been decided that linkage to hospital records would be used as a part of at least one project of this thesis. The implications of this decision will be further discussed below (see section 3.5.3 Strengths and limitations of this systematic review).

3.2 Objectives

1. To systematically identify studies on the association between depression and risk of first-ever fatal or non-fatal stroke and/ or MI
2. To investigate to what extent studies have previously explored mechanisms and mediating pathways of the depression-CVD association
3. To assess strength and weaknesses of the current evidence base

3.3 Methods

3.3.1 Search strategy

Primary studies assessing the relationship between depression and risk of first-ever fatal or non-fatal stroke or MI were identified through a comprehensive search of electronic literature databases. Based on the database's focus, EMBASE, Medline, PsycINFO, CINAHL, and Web of Science were identified as relevant databases. The search strategy was based on three search clusters: depression, stroke or MI, and association or causality. For each of the clusters, synonyms were identified through screening of relevant studies and previous published systematic reviews. A preliminary search was performed and the search strategy was discussed with a librarian. No language restrictions were made. An adapted version of the search strategy was developed for each database (Appendix 1 to Appendix 5). Since the preliminary search highlighted a large number of hits, it was decided to screen all search results backwards until 01 January 2010. This date was chosen because it was deemed likely that relevant primary studies published before January 2010 could be identified through screening of reference lists of eligible primary studies published after January 2010 and through screening of reference lists of well-conducted published systematic reviews and meta-analyses. The first search was run on all databases on 07 March 2016. To identify relevant studies that have been published

during the course of this project, an update of the systematic review was performed on 24 January 2019.

3.3.2 Study selection

3.3.2.1 Selection procedure

All retrieved articles were imported into EndNote X8 (Clarivate Analytics, 2016). Titles and abstracts were screened for potentially relevant studies. Full texts of all potentially relevant studies were read to compare the articles against *a priori* stated inclusion and exclusion criteria. If articles did not meet all inclusion criteria, the reason for exclusion was noted. Authors were contacted if eligibility could not be definitely determined after reading the full-text of articles. If authors did not respond or could not provide necessary information, the article was excluded. I tried to identify published articles of potentially eligible conference abstracts. If I could not identify a published article, I contacted the authors. If the authors did not respond or no article was published, the conference abstract was excluded.

3.3.2.2 Inclusion and exclusion criteria

Eligible participants were adults (≥ 18 years), free from stroke (when interested in stroke as outcome) or MI (when interested in MI as outcome). Participants in the exposure group either had to have a clinical diagnosis of depression or depressive symptoms. In contrast, the comparison group had to be free of clinically diagnosed depression or depressive symptoms. Additionally, studies were eligible if they reported the risk of stroke and/ or MI per unit increase on a depressive symptom rating scale. To decrease heterogeneity between studies, it was decided not to include studies that used measures of general psychological distress, some of which focus more on depression and anxiety and others capture a broader spectrum of mental health symptoms. However, since one of the projects of this thesis uses a measure of psychological distress, these studies will be discussed in the background section of the other project (see section 6.1 Background). Studies solely focusing on the effect of antidepressant or antipsychotic medication on stroke or MI risk were not eligible. The outcomes of interest were first-ever fatal or non-fatal stroke and first-ever fatal or non-fatal MI. Studies were eligible if the diagnosis was made based on self-report or

recorded diagnosis in medical records. Studies were not eligible if they defined strokes according to all CBVD ICD codes (e.g. ICD-10: I60 – 69, ICD-9: 430 – 438) or all IHD (e.g. ICD-10: I20 – I25, ICD-9: 410 – 414), except for studies in which these events were adjudicated by a medical committee. Equally, TIAs as well as coronary artery bypass grafting or percutaneous coronary interventions were not eligible as assessment of stroke and MI. Studies with participants that were selected on a condition other than depression were excluded.

3.3.3 Quality assessment

Since previous reviews pointed towards limitations in the exposure and outcome assessment and in the way potential confounding factors were considered, it was of interest to identify a quality assessment tool that allowed a detailed assessment of these factors. The Critical Appraisal Skills Programme (CASP) cohort study checklist was selected since it fulfilled the criteria mentioned above and was user friendly. The three sections of the CASP checklist allowed for an assessment of the validity of results, the results itself, and the implications and usefulness of the study (Critical Appraisal Skills Programme, 2019). One additional category was added to the checklist because the CASP checklist did not cover to what extent mediating factors/ mechanisms were discussed and taken into account in the analysis.

3.3.4 Data extraction

Data from eligible studies were extracted into pre-tested Microsoft Access tables. This form contained information on the study (e.g. author, year, country, data source), its characteristics (e.g. total sample size and size of each group, number of events, mean age, sex distribution, years of follow-up, loss to follow-up), methodology (e.g. exposure and outcome assessment, statistical method of analysis, adjustments, mediation discussed/ examined, interaction analyses, stratified analyses) and results (e.g. effect estimates for the total sample and subgroups with 95% CIs). If numbers differed between text and tables, the information of the table was extracted. If a specific value, such as number of events or individuals in the exposure group, was not reported in the study but sufficient information was provided to calculate it, the

number was calculated by the author of this dissertation. If a study compared three or more exposure groups, information of the most and least exposed groups was extracted.

3.3.5 Data synthesis

Studies which were deemed comparable were included in a random effects meta-analysis. Separate meta-analyses were performed for each of total stroke, ischaemic stroke, haemorrhagic stroke, and MI. Since the use of dichotomised and continuous exposure scales has implications on the interpretation of effect estimates, separate meta-analyses were performed for studies depending on the exposure assessment. Heterogeneity was assessed using the chi-squared test and the I^2 statistic. The I^2 statistic quantifies the proportion of variance between studies that is not explained by random variation (Higgins & Green, 2011). Subgroup analyses were conducted to investigate to what extent there were differences in the risk of stroke or MI between men and women and between studies using a measure of clinical depression and depressive symptoms. To avoid double-counting information of the same participant, only one effect estimate of each study was included in the meta-analysis. The effect estimate that was deemed most comparable to the other included studies with regard to the methodological approach was chosen. Furthermore, if more than one study used the same data source, only one of the studies was included in the meta-analysis. Studies were chosen based on the number of participants and/ or length of follow-up. The results of studies excluded from the quantitative data synthesis will be discussed and compared to the results of included studies. Meta-analyses were performed using the metagen function of the meta package in R (Schwarzer, 2007).

3.4 Results

3.4.1 Study selection

Fifty-one studies met the inclusion criteria, 36 of which investigated the relationship between depression or depressive symptoms and risk of stroke and 24 of which examined the association between depression or depressive symptoms and risk of MI (Figure 3.1).

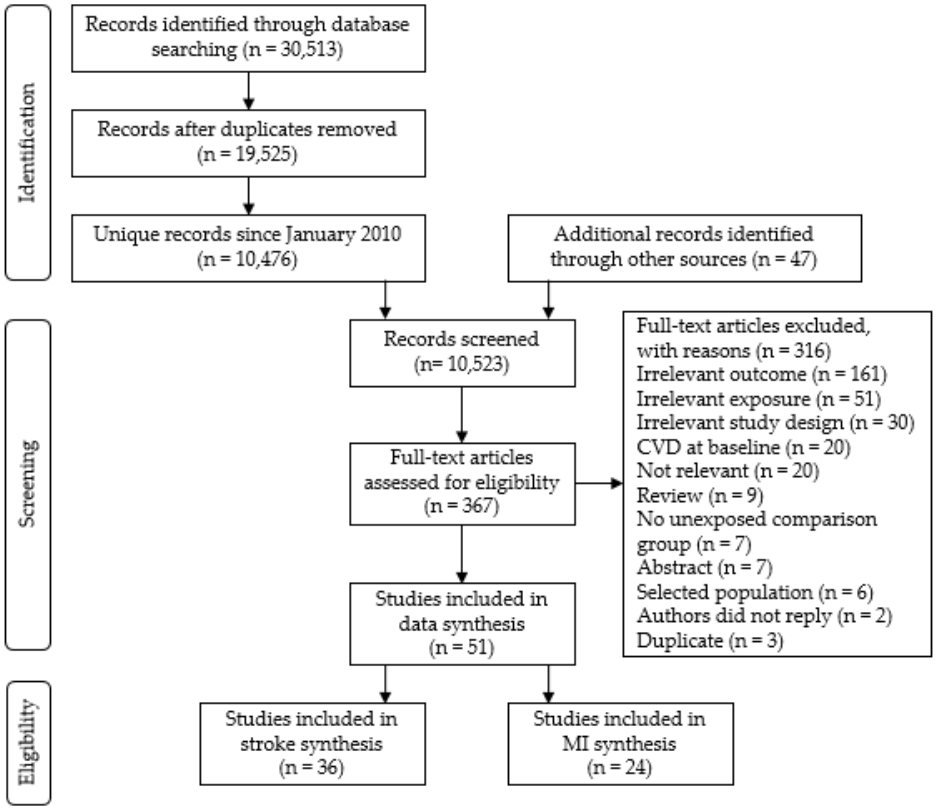


Figure 3.1: Flow-chart of study selection (adapted version of the PRISMA Flow Diagram (Moher et al, 2009))

3.4.2 Study characteristics

3.4.2.1 Eligible studies

Included studies were published from 1996 (Barefoot & Schroll, 1996; Pratt et al, 1996) to 2018 (Niles & O'Donovan, 2018). The length of follow-up ranged from 2 years (Kim et al, 2011) to 37 years (Janszky et al, 2010) and the sample sizes ranged from 574 (Gafarov et al, 2017) to 1,937,360 participants (Daskalopoulou et al, 2016). The included studies were based on 35 different data sources. Fifteen of these were based on data from the US, three from the UK, two each from Australia, Germany and the Netherlands, and one each from Denmark, Finland, Sweden, Norway, Spain, France, Russia, China, Japan, and Taiwan. One study combined data from France and Northern Ireland. Four of these data sources solely relied on routinely collected medical records (Daskalopoulou et al, 2016; Janszky et al, 2010; Mathur et al, 2016; Scherrer et al, 2010; Scherrer et al, 2011), one relied on insurance data (Chi et al, 2014;

Lin et al, 2014), and the remaining data sources obtained their data through questionnaires, interviews, and medical examinations.

3.4.2.2 Participants

Whilst the selection procedure of the exposure and comparison groups was the same within studies, the selection procedure of participants markedly differed across studies. For example, whilst all studies included participants free from stroke and/ or MI at baseline, many studies additionally excluded participants with a history of other CVD (Ariyo et al, 2000; Brown et al, 2011; Cummings et al, 2016; Daskalopoulou et al, 2016; Everson-Rose et al, 2014; Gafarov et al, 2017; Gustad et al, 2014; Joyce, 2015; Langvik & Hjemdal, 2015; Majed et al, 2012; Moise et al, 2016; Ohira et al, 2001; Péquignot et al, 2016; Péquignot et al, 2013; Pössel et al, 2015; Pratt et al, 1996; Scherrer et al, 2010; Scherrer et al, 2011; Stürmer et al, 2006; Sun et al, 2016; Wassertheil-Smoller et al, 2004), a history of other CVD or cardiac procedures (Stewart et al, 2016), or any known medical conditions at baseline (Kubzansky et al, 2006; Sesso et al, 1998) thereby creating samples of healthy participants at baseline. Four studies compared results after exclusion of participants with a history of cardiovascular events other than stroke and/ or MI at baseline to results with exclusion of participants with a history of stroke and/ or MI only (Barefoot & Schroll, 1996; Everson et al, 1998; Marijnissen et al, 2014; Wouts et al, 2008). One study excluded participants with any CVD at baseline (Mejía-Lancheros et al, 2014) but selected participants with an above average risk of developing events during follow-up by selecting participants with known cardiovascular risk factors at baseline. Another notable selection criterion was observed in the study by Arbelaez et al (2007) in which participants with a haemorrhagic stroke during follow-up were excluded. The authors justified this by the lack of association between depression and haemorrhagic stroke in the sample.

There were differences in the age and sex distribution of participants across studies. Forty-two studies enrolled both male and female participants, five studies enrolled men only (Janszky et al, 2010; Kubzansky et al, 2006; Majed et al, 2012; Pössel et al, 2015; Sesso et al, 1998) and three studies included women only (Jackson & Mishra,

2013; Pan et al, 2011a; Wassertheil-Smoller et al, 2004). Of the 39 studies that reported their participant's mean age at baseline, one study reported a mean age below 39 years, four studies reported a mean age from 40 – 49 years, eight studies reported a mean age from 50 – 59 years, 19 studies reported a mean age from 60 – 69 years, and seven studies reported a mean age over 70 years. Of the twelve studies that did not report their participants' mean age, four studies reported a median age of participants. One study reported a median age from 40 – 49 years, one study reported a median age from 50 – 59 years, and two studies reported a median age over 70 years. Eight studies reported the range of their participants at baseline, of which one study each included participants aged 18 – 20 years, aged 50 – 60 years, aged at least 65 years, and aged at least 75 years. Furthermore, three studies included participants aged at least 18 years at baseline.

Participant characteristics of individuals in the exposure group were compared to characteristics of those in the comparison group in 23 studies (Brown et al, 2011; Cummings et al, 2016; Daskalopoulou et al, 2016; Everson et al, 1998; Glymour et al, 2010; Jackson & Mishra, 2013; Janszky et al, 2010; Joyce, 2015; Köhler et al, 2013; Larson et al, 2001; Lin et al, 2014; Majed et al, 2012; Mathur et al, 2016; Mejía-Lancheros et al, 2014; Moise et al, 2016; O'Brien et al, 2015; Pan et al, 2011a; Péquignot et al, 2013; Pratt et al, 1996; Scherrer et al, 2010; Sesso et al, 1998; Stürmer et al, 2006; Sun et al, 2016). All of these studies reported higher proportions of women among the exposure group than among the comparison group. In ten of the studies participants in the exposure group were younger, in six studies participants in both groups were of similar age, in six studies they were older, and two studies did not report differences in age. Fifteen studies reported a lower SES among the exposed group, three studies reported higher educational attainment, one study reported no differences between the SES of the groups, and one study reported higher unemployment despite higher educational attainment among the exposure group. Furthermore, participants in the exposure group were less likely to be married, and generally had a worse cardiovascular risk profile and more comorbidities at baseline.

A notable exception is the difference in alcohol intake. Whilst seven studies reported a higher alcohol intake or more alcohol abuse/ dependence among the group with depression, eight studies reported a lower intake of alcohol among the exposure group.

3.4.2.3 Exposure assessment

The included studies varied in their way of assessing the exposure. Thirty-four studies assessed depression based on depressive symptom rating scales, 11 studies used a measure of clinical depression, three studies combined information obtained through rating scales and diagnostic interviews, two studies combined information on clinical depression and antidepressant use, and one study combined information obtained from a rating scale and/ or clinical depression and/ or antidepressant use (Table 3.1).

Differences were not only present in the use of different rating scales but also in the way different studies used the same rating scale. Nine different depressive symptom rating scales were used, of which the CES-D was the scale that was used most often (22 studies). Whilst some studies used the recommended cut-off score of at least 16 to define the presence of depressive symptoms, other studies formed quartiles, used the CES-D scale as continuous exposure scale, or used two exposure waves to define constant low, recent onset, recently remitted, and stable high depressive symptoms. Similarly, different data sources were used to define clinical depression. Whilst some studies defined clinical depression based on ICD codes (versions 8 or 9), others used Read codes, the Diagnostic Interview Schedule (DIS), or self-report of (doctor) diagnoses.

Whilst most studies assessed depressive symptoms or clinical depression at baseline only, twelve studies investigated the influence of symptom severity, timing of symptoms, and symptom recurrence on effect estimates. Five of these studies assessed the risk of stroke and/ or MI using time-updated information on depressive symptom in statistical models (Everson et al, 1998; Glymour et al, 2010; Jackson & Mishra, 2013; Moise et al, 2016; Yan et al, 2013). Two studies calculated separate effect

estimates of the risk of stroke and/ or MI among individuals with current depression and individuals with a history of depression (Daskalopoulou et al, 2016; Pan et al, 2011a). Two studies assessed depressive symptom scores at multiple waves and calculated a mean score (Ariyo et al, 2000; Wouts et al, 2008). Wouts et al (2008) additionally investigated the risk of stroke/MI dependent on the proportion of observations with clinical depression or depressive symptoms. Péquignot et al (2016) assessed depressive symptoms at multiple waves and modelled the risk of stroke as a function of the count of occasions with high depressive symptoms. Gilsanz et al (2017); Gilsanz et al (2015) investigated the risk of stroke dependent on changes in depressive symptoms. Changes in depressive symptoms were defined according to two exposure waves preceding the outcome wave. Participants were allocated to a group with stable high depressive symptoms, recent onset depressive symptoms, recently remitted depressive symptoms, and stable low depressive symptoms depending on whether they had depressive symptoms (CES-D score of at least 16) in both, neither, the first exposure wave only or the second exposure wave only.

3.4.2.4 Outcome assessment

Thirty-six studies reported the association between depression or depressive symptoms and subsequent stroke. Thirty-two of these studies investigated the association between depression or depressive symptoms and total stroke, ten studies explored the relationship with ischaemic stroke, six studies assessed the risk of haemorrhagic stroke, and one study looked at the relationship between depression or depressive symptoms and risk of uncertain type of stroke. Twenty-four studies reported the association between depression or depressive symptoms and subsequent MI.

Studies varied greatly with regard to what data sources were used to determine events during follow-up (Appendix Table A 1). Thirty-eight studies had access to at least one source of routinely collected medical records including death records. Twenty-four studies had access to hospital records of participants. Furthermore, 16 studies obtained information on events in primary care. Eleven of these studies relied

on confirmation through the participant's GP whereas five studies had access to routinely collected primary care records. Thirty studies used death records to determine the date and cause of death. Eight studies additionally used coroner reports, autopsy reports or obituaries to identify stroke and/ or MI events. Seven studies solely relied on self-report of participants or proxies, six of which were based on data from the Health and Retirement study. Two studies relied on national health insurance records but did not further specify what type of records were collected. One study did not give any information on data sources used to establish events during follow-up (Sesso et al, 1998).

Table 3.1: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 1 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Arbelaez et al (2007)	Cardiovascular Health Study, US	Random sample of Medicare-eligible, non-institutionalised adults, aged ≥ 65 years, residing in one of four US counties, recruited 1989 – 1991. A second enrolment of African-Americans completed 1992 – 1993. Participants being wheelchair bound in the home, undergoing radiation or chemotherapy for cancer and with a haemorrhagic stroke during follow-up were excluded.	5,525	11†	72.7 (5.6)	41.8	Depressive symptoms (CES-D)	Ischaemic stroke
Ariyo et al (2000)	Cardiovascular Health Study, US	Random sample of Medicare-eligible, non-institutionalised adults, aged ≥ 65 years and an additional ethnic minority cohort, which was recruited later.	4,493	≤ 6	72.4	39.1	Depressive symptoms (CES-D)	MI
Avendano et al (2006)	EPESE study, US	Men and women, aged ≥ 65 years, from a sample stratified by type of residence. Participants with missing baseline demographics were excluded.	2,250	≤ 12	NR	NR	Depressive symptoms (CES-D)	Total stroke
Barefoot & Schroll (1996)	Residents of Glostrup, Denmark	Residents of Glostrup, Denmark, born in 1914, were invited to participate in 1964 and 1974. Participants who did not participate in at least one psychological examination in 1964 or 1974 were excluded.	675	≤ 27	NR	52.8	Depressive symptoms (MMPI OBD subscale)	MI
Bos et al (2008)	Third Rotterdam Study Survey, Netherlands	Adults, aged ≥ 61 years, living in one district of Rotterdam 1990 – 1993. Participants who refused the interview, had incomplete information on the CES-D, or were physically unable to attend were excluded.	4,424	6 – 8	71.9†	39.8	Depressive symptoms (CES-D) (and clinical depression (Present State Examination)	Total stroke, ischaemic stroke

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 2 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Brown et al (2011)	An urban primary care practice, US	Primary care patients, aged ≥ 60 years, who attended scheduled appointments and were screened for depression. Patients unable to answer more than four CES-D questions, with severe cognitive impairment, hearing impairment, non-English speakers, imprisoned or institutionalised patients were excluded.	2,728	13 – 16	67.5	28.6	Depressive symptoms (CES-D)	MI
Brunner et al (2014)	Whitehall II study, UK	London based civil servants, not including blue-collar workers, aged 35 – 55 years at phase 1 (1985 – 1988). Phase 7 was used as baseline for this analysis (2003 – 2004).	5,717	5	61.0 (6.0)	70.3	Depressive symptoms (CES-D)	Total stroke
Chi et al (2014)	National Health Insurance Database, Taiwan	Participants with a depression diagnosis between 2002 and 2005 were compared with age, gender, index year-matched participants with no history of depression before 2010 (end of follow-up).	132,090	6	< 45: 48.3%, 45 – 64: 32.9%, ≥ 65: 18.8%	37	Clinical depression (ICD-9-CM: 296.2, 296.3, 300.4, 311)	MI
Cummings et al (2016)	REGARDS study, US	Adults with/ without diabetes, aged > 45 years, recruited through a population-based sampling method between January 2003 and October 2007. Participants with missing diabetes status or incomplete follow-up were excluded.	22,003	5.95†	63.9	42	Depressive symptoms (CES-D)	Total stroke MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 3 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Daskalopoulou et al (2016)	CALIBER programme, UK	Patients registered with 225 general practices, aged ≥ 30 years, with at least one year of pre-study follow-up within a Clinical Practice Research Datalink general practice.	1,937,360	6.9 (3.2 – 10.5)†	48.3 (15.6)‡	33	History/ new onset clinical depression (CPRD) and/ or antidepressant use (CPRD)	Ischaemic stroke, intracerebral & subarachnoid haemorrhage, MI
Everson et al (1998)	Alameda County Study, US	Representative sample of the adult, non- institutionalised population of Alameda County, California, aged 17 – 94 years. Data from the 1965, 1974, and 1983 were used in this analysis. Participants with missing or incomplete data on the depression measure, or missing data on covariates were excluded.	6,676	≤ 29	43.4 (15.9)	45.8	Depressive symptoms (HPL)	Total stroke
Everson-Rose et al (2014)	Multi-Ethnic Study of Athero-sclerosis, US	Adults, aged 45 – 84 years, enrolled at six field sites between July 2000 and August 2002, without missing data on all psychosocial measures and demographic variables.	6,643	8.5 (0.0 –10.9)§	62.1 (10.2)	47.1	Depressive symptoms (CES-D)	Total stroke/ ischaemic stroke
Gafarov et al (2017)	WHO MONICA, Russia	Nationally representative sample of adults, aged 25 – 64 years, from one district in Novosibirsk. The sample was based on electoral lists of Russian citizens.	574	≤ 16	Men: 44.3 (0.4) Women: 45.4 (0.4)		Depressive symptoms (MOPSY)	Total stroke MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 4 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Gilsanz et al (2015)	Health and Retirement Study, US	Nationally representative sample of adults and their spouses, enrolled from 1992 – 1998. This analysis included non-institutionalised participants, aged ≥ 50 in 1998, without missing values on baseline depression score or baseline covariates.	16,178	8.9	65.7 (9.7)	41.5	Changes in depressive symptoms (CES-D)	Total stroke
Gilsanz et al (2017)	Cardiovascular Health Study, US	Random sample of Medicare-eligible adults, aged ≥ 65, recruited in four field sites from 1989 – 1990. An additional sample of African-Americans was recruited from 1992 – 1993. Participants with missing baseline covariates, missing depression scores during either of the first two exposure waves, or a missing interview before the end of the first exposure period were excluded.	4,319	9.0 (2.0 – 10.9)§	72.4 (5.2)	41.6	Changes in depressive symptoms (CES-D)	Total stroke/ ischaemic stroke/ haemorrhagic stroke
Glymour et al (2010)	Health and Retirement Study, US	Nationally representative sample of adults, aged ≥ 50, and their spouses, enrolled from 1992 – 1998. 1996 was considered baseline for this analysis. Participants with missing depression or memory impairment information at baseline, and missing covariate or follow-up information were excluded.	19,087	8.1	65.8	41	Depressive symptoms (CES-D)	Total stroke

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 5 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Glymour et al (2012)	Health and Retirement Study, US	Nationally representative sample of adults, aged ≥ 50, and their spouses, enrolled from 1992 – 1998. 1996 was considered baseline for this analysis. Participants with missing baseline depression, baseline covariates, failure date, or follow-up information were excluded.	18,648	8.2	65.8	40.9	Depressive symptoms (CES-D)	Total stroke
Gustad et al (2014)	Nord – Trøndelag Health Study, Norway	All adult citizens living in Nord-Trøndelag County were invited, those who responded and had a valid HADS-D response (responded to ≥ 5 items) were included.	57,819	11.4 (2.9)	47.7 (16.3)	45.8	Depressive symptoms (HADS-D)	MI
Henderson et al (2013)	Chicago Health and Aging Project, US	Community-dwelling black and non-Hispanic white adults, aged ≥ 65 years at first cycle (1993 – 1996), sampled from three adjacent neighbourhoods within the south side of Chicago, without missing data on psychosocial factors, demographic characteristics, and vital status. Participants who were actively involved in a health maintenance organisation during the time of follow-up were excluded. The second cycle (1997 – 1999) was used as baseline.	2,539	6.0 (3.4)	77.1 (6.3)	38.2	Depressive symptoms (CES-D)	Haemorrhagic stroke

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 6 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Jackson & Mishra (2013)	Australian Longitudinal Study on Women's Health, Australia	Women from the 1946 – 1951 cohort aged 47 – 52 years, randomly selected from the Medicare database. The 1998 survey was used as baseline. Women who did not complete stroke questions at any point or had no data on depression and stroke occurrence for at least one survey period were excluded.	10,547	≤ 12	52.5 (1.5)	0	Depressive symptoms (CES-D) and/ or antidepressant use (self-report) and/ or clinical depression (self-report)	Total stroke
Janszky et al (2010)	Conscription surveys and health registers, Sweden	Nationwide survey of men, aged 18 – 20, born 1949 – 1951, who were conscripted for compulsory military service in 1969 and 1970.	49,321	37 (1)	NR	100	Clinical Depression (ICD-8: 296 or 300.4)	MI
Joyce (2015)	45 and Up Study, Australia	Residents of New South Wales, Australia, aged ≥ 45, randomly sampled from the Australian Medicare database, oversampling of adults aged ≥ 80 and residents of non-urban areas. Participants with cancer, where missing data categories would result in small cell sizes, and those who completed versions of baseline survey with no differentiation between depression and anxiety were excluded.	143,815	2.4	60 (10)	41.9	Clinical Depression (self-report)	MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 7 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Kim et al (2011)	Health and Retirement Study, US	Nationally representative sample of adults, aged ≥ 50, and their spouses, enrolled from 1992 – 1998. The eighth wave (2006) was used as baseline. Results are based on complete case analysis.	6,044	≤ 2	68.5 (9.6)	42.1	Depressive symptoms (CES-D)	Total stroke
Kim et al (2013)	Health and Retirement Study, US	Nationally representative sample of adults, aged ≥ 50, and their spouses, enrolled from 1992 – 1998 who completed a self-reported psychological questionnaire in the eighth wave (baseline of this analysis).	6,739	≤ 4	68.8 (9.8)	41.6	Depressive symptoms (CES-D)	Total stroke
Köhler et al (2013)	German Study on Ageing, Cognition, Dementia in Primary Care, Germany	Primary care patients, aged ≥ 75 years, recruited from 138 general practices in six study centres. Participants with dementia diagnosis, missing data, without one GP contact in the past 12 months, people residing in nursing homes, those with fatal illness leading to death within three months, insufficiency of German language, deafness and blindness, incapacity to give informed consent, and those not a regular patient of the GP were excluded.	2,854	≤ 6	75 – 79: 54.0%, 80 – 84: 37.0%, ≥ 85: 9.0%	44	Depressive symptoms (GDS)	Total stroke
Kubzansky et al (2006)	Normative Aging Study, US	Community-dwelling men from the greater Boston area, aged 40 – 90 in 1986, free of known chronic medical conditions in 1961. The study was established by the Veterans Administrations.	1,306	10.9 (3.3)	61 (8.3)	100	Depressive symptoms (factor analysis identified iso-depression based on MMPI-2)	MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 8 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Langvik & Hjemdal (2015)	Nord – Trøndelag Health Study, Norway	Adult citizens living in Nord-Trøndelag County, participating in HUNT 2 (baseline of this analysis).	28,859	5 – 8	47.0 (13.0)	43.1	Depressive symptoms (HADS-D)	MI
Larson et al (2001)	Baltimore Epidemiologic Catchment Area Study, US	Adults, aged ≥ 18 years, selected through probability sampling within the east Baltimore area from 1980 – 1983, with oversampling of those aged ≥ 65 years. Participants without follow-up information, who did not know whether they had a stroke before baseline, and participants reporting mania were excluded.	1,703	≤ 13	18 – 29: 33.9%, 30 – 44: 30.4%, 45 – 54: 10.9%, 55 – 64: 13.6%, ≥ 65: 11.3%	37.1	Clinical Depression (DIS)	Total stroke
Lin et al (2014)	National Health Insurance Research Database, Taiwan	Participants were selected from a random sample of 1,000,000 insured adults, aged ≥ 18, released by the National Health Research Institute in 2009. Participants with a depression diagnosis in the period of 2000 – 2002 were compared with age-, sex-, and index year-matched insured adults without depression in the same period. Participants with missing information on age or sex were excluded.	54,355	10	18 – 44: 45.8%, 45 – 64: 33.8%, 65 – 74: 13.4%, ≥ 75: 7.0%	38.4	Clinical depression (ICD-9-CM: 296.2, 296.3, 300.4, 311)	MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 9 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Majed et al (2012)	PRIME study, France and Northern Ireland	Men, aged 48 – 64 years, recruited from four WHO MONICA centres in France or Northern Ireland. Participants with several missing covariates or missing depressive symptom score were excluded.	8,746	10	55	100	Depressive symptoms (CES-D)	Total stroke/ ischaemic stroke
Marijnissen et al (2014)	Longitudinal Aging Study Amsterdam, Netherlands	Participants with/ without cardiac disease, aged ≥ 55 years, with available baseline data on depressive symptoms, stroke, and neuroticism.	2,050	≤ 9	69.3 (8.5)	49.0	Depressive symptoms (CES-D)	Total stroke
Mathur et al (2016)	East London Primary Care Database, UK	Adults, aged ≥ 30 years, registered in the database in March 2015 (end of follow-up). The database comprises the electronic health records of 950,000 individuals registered with 141 general practices across four east London boroughs.	524,952	10	35.9 (13.9)	52.8	Clinical Depression (Read codes)	Total stroke MI
Mejía-Lancheros et al (2014)	Prevention with Mediterranean diet study, Spain	Adults at high CVD risk, aged 55 – 80 years, without severe medical condition (digestive disease with fat intolerance, advanced malignancy, major neurological, psychiatric or endocrine disease), with complete data on psychosocial risk factors (depression, educational level, social support).	7,263	4.8 (2.8 – 5.8)†	67.0 (6.2)	42.5	Clinical Depression (self-report)	Total stroke MI
Moise et al (2016)	REGARDS study, US	African-American and white adults, aged ≥ 45 years, recruited by mail using commercially available lists of US residents, without missing depressive symptom score at baseline.	22,666	6.9†	63.9 (9.3)	41.2	Depressive symptoms (CES-D)	Total stroke

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 10 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Niles & O'Donovan (2018)	Health and Retirement Study, US	Nationally representative sample of adults, aged ≥ 50, enrolled either in 2006 or 2008 with at least one available anxiety and depression measure. Results are based on complete case analysis.	10,843	4	68.0 (10.5)	40.9	Depressive symptoms (CES-D)	Total stroke
O'Brien et al (2015)	Jackson Heart Study, US	African-American participants, aged 21 – 94 years, from four populations: Community volunteers from the Jackson metropolitan area (30%), randomly selected residents of Jackson (17%), participants in the Jackson site of the Atherosclerosis Risk in Communities cohort study (22%), and family members of Jackson Heart Study participants (31%). Participants were excluded if they responded to less than 16 CES-D questions.	3,309	8†	~53*	34.6	Depressive symptoms (CES-D)	Total stroke
Ohira et al (2001)	Circulatory Risk in Communities Study, Japan	Adults, aged 40 – 78 years, who were examined in a rural community (Kyowa, Ibaraki preference) in October 1985.	884	10.3	56.5	35.4	Depressive symptoms (Zung SDS)	Total stroke, Ischaemic & haemorrhagic stroke

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 11 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Pan et al (2011a)	Nurses' Health Study, US	Female registered nurses, aged 54 – 79 in year 2000 (baseline of this analysis). Participants with no information on psychological distress, depression diagnosis, or antidepressant use and with missing values for covariates were excluded.	80,574	≤ 6	66	0	Current/ past depression diagnosis (self-report) and/ or antidepressant use (self-report)	Total stroke, ischaemic & haemorrhagic stroke, stroke of unknown type
Péquignot et al (2013)	The Three – City Study, France	Non-institutionalised participants, aged ≥ 65 years, randomly selected from electoral rolls of three large French cities between March 1999 and March 2001. Participants with dementia or no CES-D score at baseline were excluded.	7,308	5.3 (4.8 – 5.6)†	73 (69 – 77)†	36.5	Depressive symptoms (CES-D)	Total stroke MI
Péquignot et al (2016)	The Three – City Study, France	Non-institutionalised participants, aged ≥ 65 years, randomly selected from electoral rolls of three large French cities between March 1999 and March 2001. Participants with dementia or no CES-D score at baseline were excluded.	7,313	8.4 (2.3)†	73.8 (5.4)	36.6	Depressive symptoms (CES-D)	Total stroke
Pössel et al (2015)	Kuopio Ischaemic Heart Disease Study, Finland	Men, aged 42 – 61 years, from the Kuopio region in Eastern Finland, enrolled between March 1984 and December 1989.	2,005	15.2 (12.3 – 18.0)	52.5 (5.3)	100	Depressive symptoms (HPL)	MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 12 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Pratt et al (1996)	Baltimore Epidemiologic Catchment Area Study, US	Area-probability sample of household residents of east Baltimore, aged ≥ 18 years, interviewed from 1980 – 1983. Participants who did not know whether they had ever had a heart attack were excluded from this analysis.	1,551	≤ 13	18 – 29: 35.9% 30 – 44: 31.5% 45 – 54: 10.7% 55 – 64: 12.9% ≥ 65: 9.0%	37.6	Clinical Depression (DIS)	MI
Scherrer et al (2010)	Veterans Administration electronic medical records, US	Veterans, aged 25 – 80, free from affective psychoses, manic disorders, bipolar disorder, and hypothyroidism, with at least one outpatient visit in the two years before baseline.	355,999	≤ 7	55.7 (13.2)	88.2	Clinical diagnosis (ICD-9-CM codes: 296.2, 296.3, 300.4 or 311)	MI
Scherrer et al (2011)	Veterans Administration electronic medical records, US	Veterans without diabetes, aged 25 – 80 years, with at least one outpatient visit recorded in the two years before baseline. Participants with a psychotic or bipolar disorder, dysthymia diagnosis without depression, less than 12 weeks of follow-up and a MI in the first month of follow-up were excluded.	292,317	≤ 7	55.6 (13.2)	88.3	Clinical Diagnosis (ICD-9-CM codes: 296.2, 296.3, or 311)	MI
Sesso et al (1998)	Normative Aging Study, US	Community- dwelling men from the greater Boston area, aged 40 – 90 years at baseline, free of known chronic medical conditions at the start of the study (1961). The study was established by the Veterans Administrations.	1,305	7.0 (2.3)	61.8 (8.3)	100	Depressive symptoms (MMPI-2 D, MMPI-2 Dep, SCL-90)	MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 13 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Stewart et al (2016)	IMPACT study (Indiana site), US	Primary care patients, aged ≥ 60 years, attending two large urban primary care clinics in a safety-net health care system in Indianapolis, were approached for screening during routine visits. Patients for whom the IMPACT trial data could not be matched to other sources (e.g. Medicare claims) and patients who were randomised in the IMPACT trial were excluded.	2,041	8.3†	68.5 (6.9)	27.4	Depressive symptoms (modified PRIME-MD questionnaire)	Total stroke MI
Stürmer et al (2006)	Population registry of Heidelberg, Germany	Representative sample of adults, aged 40 – 65 years, from the population registry of Heidelberg, Germany. Participants with missing information (e.g. unknown address) on disease status were excluded.	4,267	8.5†	53.4 (7.1)	48.5	Depressive symptoms (factor analysis of personality scales)	Total stroke MI
Sun et al (2016)	China Kadoorie Biobank cohort, China	Men and women aged 30 – 79 years, enrolled 2004 – 2008, from ten diverse urban and rural areas in China. Participants with a history of cancer, and one participant with implausible censoring date were excluded.	487,377	7.2†	51	40.9	Clinical depression (CIDI-SF)	Total stroke/ ischaemic stroke, haemorrhagic stroke
Wassertheil-Smoller et al (2004)	Women's Health Initiative Observational Study, US	Postmenopausal women, aged 50 – 79 years, recruited at four clinical centres in the US, mostly through mass mailings to age-eligible women from large mailing lists. Women who participated in other RCTs, with predicted survival of less than 3 years, alcoholism, drug dependency, diagnosed mental illness, dementia, or other conditions making women unable to participate were excluded.	73,098	4.1 (1.2)	63.8	0	Depressive symptoms (CES-D) and responded "yes" to 2 DIS items	Total stroke MI

Table 3.1 continued: Characteristics of studies on the association between clinical depression or depressive symptoms and risk of stroke and/ or myocardial infarction (page 14 of 14)

Authors, Year	Data source	Population characteristics	Sample size	Years of follow-up	Mean age (SD)	Male (%)	Ascertainment depression	Outcome(s)*
Wouts et al (2008)	Longitudinal Aging Study Amsterdam, Netherlands	Adults with/ without cardiac diseases, aged 55 – 85 years, randomly selected from an age- and sex-stratified sample drawn from population registers of 11 municipalities. Participants in whom depressive symptoms or stroke at baseline were not evaluated were excluded.	2,965	7.7 (3.1)	70.5 (8.7)	47.9	Depressive symptoms (CES-D) and/ or clinical depression (DIS)	Total stroke
Yan et al (2013)	Cardiovascular Health Study, US	White and African-American adults from four US counties, aged ≥ 65 years, recruited from 1992 – 1993.	4,619	≤ 11.5	72.8	40.9	Depressive symptoms (CES-D)	Ischaemic stroke

* More detailed information on data sources used to establish events during follow-up is presented in Appendix Table A 1
† Median (IQR, if provided)
‡ Based on participants with a history of depression at baseline
§ Median (range)
CALIBER: Cardiovascular research using Linked Bespoke studies and Electronic Health Records programme, CES-D: Centre for Epidemiologic Studies Depression Scale, CIDI-SF: Composite International Diagnostic Interview – Short Form, CRPD: Clinical Research Practice Datalink, CVD: Cardiovascular diseases, DIS: Diagnostic Interview Scale, EPESE: Established populations for epidemiologic studies of the elderly, GDS: Geriatric Depression Scale, GP: General Practitioner, HADS-D: Hospital Anxiety and Depression Scale - Depression subscale, HPL: Human Population Laboratory Depression scale, ICD: International Classification of Diseases, IMPACT: Improving Mood-Promoting Access to Collaborative Treatment, MI: Myocardial infarction, MMPI OBD: Minnesota Multiphasic Personality Inventory obvious depression subscale, MMPI-2: Second version of the Minnesota Multiphasic Personality Inventory, MMPI-2 D: Second version of the Minnesota Multiphasic Personality Inventory - Depression subscale in keeping with Butcher et al (1989), MMPI-2 Dep: Second version of the Minnesota Multiphasic Personality Inventory - Depression subscale in keeping with Hathaway et al (1951), MOPSY: MONICA psychosocial Interview Depression scale, NR: Not reported, PRIME-MD: Primary Care Evaluation of Mental Disorders, RCT: Randomised controlled trial, REGARDS: REasons for Geographic And Racial Differences in Stroke study, SCL-90: Symptom Checklist-90, SD: Standard deviation, UK: United Kingdom, US: United States of America, WHO: World Health Organisation, Zung SDS: Zung Self-Rating Depression scale

3.4.2.5 Conditioning on covariates

All studies addressed potential confounding and/ or mediating factors to some extent, either by adjusting for covariates in statistical models, by restricting the sample, by reporting stratified results, or by applying inverse probability weights (Table 3.2). All studies reported effect estimates after conditioning on age and all but one study conditioned on sex. Thirty-one studies conditioned on SES. Whilst most studies used a measure of educational attainment to condition on SES, others used a measure of occupation, employment status, income, wealth, area-based deprivation, and homeownership. Other sociodemographic factors that were taken into account by some but not all studies were race/ ethnicity, marital status, region/ urbanisation, and social support. Forty studies conditioned on at least one lifestyle factor with smoking being the most common factor conditioned on (40 studies), followed by at least one measure of body weight (32 studies). However, only eleven studies conditioned on each of smoking, physical activity, alcohol intake, and a measure of weight (e.g. body mass index (BMI)) of which two studies additionally conditioned on at least one dietary factor. Forty-seven studies conditioned on at least one comorbidity at baseline. Of these, 42 studies conditioned on blood pressure/ hypertension, 40 studies on diabetes mellitus, 24 studies on cholesterol levels, and 41 studies on CVD (other than stroke or MI at baseline). Thirteen studies conditioned on psychopathology other than depression at baseline. Fourteen studies conditioned on at least one medication, ten studies on at least one biological factor, nine studies on at least one factor related to the family of participants, and seven studies on at least one study-design related factor.

Whilst the majority of studies conditioned on covariates at baseline only, six studies conditioned on time-updated information of covariates during follow-up. Three of these studies adjusted for time-varying covariate information in Cox proportional hazards models (Everson et al, 1998; Pan et al, 2011a; Wouts et al, 2008), one study included time-updated information in a generalised estimating equation (GEE) regression model (Jackson & Mishra, 2013), and two studies applied stabilised inverse probability weights (Gilsanz et al, 2017; Gilsanz et al, 2015). The weights were applied

in such a way that depressive symptoms were no longer associated with time-updated covariates, survival, and study dropout.

Whilst most studies discussed the role of different pathways or mediating factors in the relationship between depression and cardiovascular events in the background and discussion sections, only 12 studies took the role of mediating factors into account in their analysis. Four of these studies hypothesised that some of the covariates might be on the causal pathways but adjusted for these factors using the same approach as for hypothesised confounding factors (Majed et al, 2012; Moise et al, 2016; O'Brien et al, 2015; Péquignot et al, 2016). Janszky et al (2010), Glymour et al (2010), Glymour et al (2012) and Gustad et al (2014) hypothesised that some of the covariates were on the causal pathway and reported effect estimates with and without inclusion of these covariates. Furthermore, Glymour et al (2010; 2012) highlighted that the effect estimates after inclusion of potential mediating factors should not be interpreted causally. Wassertheil-Smoller et al (2004) identified some of the covariates as mediators and stated that these variables were not included in the multivariate model since they were assumed to be on the causal pathway. Gilsanz et al (2015; 2017) considered the timing of exposure, covariates, and outcome and stated that inverse probability weights were applied to control for time-updated covariates since they might act as confounders as well as mediators. Arbelaez et al (2007) aimed to investigate the potential role of inflammation as a mediator. For that, the authors investigated to what extent the effect estimate was altered when results were first adjusted and then stratified for C-reactive protein tertiles.

Table 3.2: Descriptive summary of covariates in included studies (page 1 of 5)

	Sociodemographic factors											Lifestyle factors					Comorbidities/ impairment									Biological factors			Medication					Study factors		Familial factors		
Author	Age	Sex	Marital status	Race/ethnicity	Region/ urbanisation	Education	Occupation/ employment status	Income/ wealth	Area-based deprivation	Homeownership	Social support/ social deprivation	Smoking	Physical inactivity	Alcohol use	BMI/obesity/weight/ waist circumference	Diet/ dietary supplements	CVD (other than stroke/MI)	Blood pressure/ hypertension	Diabetes (mellitus)	Cholesterol/ lipid levels	Cancer	Psychopathology (other than depression)	Physical impairment	Cognitive impairment/ function	Height	C-reactive protein	Menopausal status/ HRT use	Medication adherence	Antidepressants/ antipsychotics	Antihypertensive medication	Lipid-/ cholesterols-lowering drugs	Aspirin/ antithrombotic drugs	Study site	Year of enrolment	Mother's/ father's SES	Family history of CVD/ CHD	**Others***
Arbelaez et al (2007)	✓	✓	✓	✓		✓	✓	✓				✓			✓		✓	✓	✓	✓	✓†					✓											✓
Ariyo et al (2000)	✓	✓	✓	✓		✓						✓	✓	✓			✓†	✓	✓	✓																
Avendano et al (2006)	✓	✓		✓		✓		✓																													
Barefoot & Schroll (1996)	✓	✓										✓	✓					✓		✓																✓
Bos et al (2008)	✓	✓										✓					✓	✓	✓										✓	✓	✓	✓				✓
Brown et al (2011)	✓	✓		✓								✓			✓		✓†	✓	✓	✓																
Brunner et al (2014)	✓	✓		✓																																
Chi et al (2014)	✓	✓			✓			✓																										✓		✓
Cummings et al (2016)	✓	✓		✓	✓	✓		✓				✓			✓		✓	✓	✓‡	✓						✓					✓					✓
Daskalopoulou et al (2016)	✓	✓		(✓)					✓			✓		(✓)	(✓)		✓†	✓	✓	✓																✓
Everson et al (1998)	✓	✓		✓		✓§						✓§		✓§	✓§		(✓†)	✓§	✓§																	
Everson-Rose et al (2014)	✓	✓		✓		✓						✓	✓	✓	✓		✓†	✓	✓	✓					✓					✓			✓			
Gafarov et al (2017)	✓	✓‡	✓			✓	✓										✓†	✓†	✓†																	

Table 3.2 continued: Summary description of covariates in included studies (page 2 of 5)

Author	Sociodemographic factors												Lifestyle factors					Comorbidities/ impairment								Biological factors			Medication					Study factors		Familial factors		Others*
	Age	Sex	Marital status	Race/ ethnicity	Region/ urbanisation	Education	Occupation/ employment status	Income/ wealth	Area-based deprivation	Homeownership	Social support/ social deprivation	Smoking	Physical inactivity	Alcohol use	BMI/obesity/weight/ waist circumference	Diet/ dietary supplements	CVD (other than stroke/MI)	Blood pressure/ hypertension	Diabetes (mellitus)	Cholesterol/ lipid levels	Cancer	Psychopathology (other than depression)	Physical impairment	Cognitive impairment/ function	Height	C-reactive protein	Menopausal status/ HRT use	Medication adherence	Antidepressants/ antipsychotics	Antihypertensive medication	Lipid-/ cholesterols-lowering drugs	Aspirin/ antithrombotic drugs	Study site	Year of enrolment	Mother's/ father's SES	Family history of CVD/ CHD	
Gilsanz et al (2015)	✓***	✓**	✓	✓**		✓**		✓				+✓					✓	✓	✓										‡	✓						✓
Gilsanz et al (2017)	✓ **	✓ **	✓	✓**		✓**		✓				(✓) (‡‡)	(✓) (‡‡)	(✓) (‡‡)	(✓) (‡‡)		(✓) (‡‡)	(✓) (‡‡)	(✓) (‡‡)					(✓) (‡‡)						(‡‡ ✓)		✓ **	✓ **			✓
Glymour et al (2010)	✓	✓	(✓)	✓		(✓)		(✓)				(✓)	(✓)	(✓)	(✓)		(✓)	(✓)	(✓)					✓										✓			✓
Glymour et al (2012)	✓	✓	✓			✓		✓				✓		✓	✓		✓	✓	✓															✓			✓
Gustad et al (2014)	✓	✓	✓			✓						✓	✓	✓	✓		✓	✓	✓	✓																	✓
Henderson et al (2013)	✓	✓		✓		✓						✓	✓		✓		✓	✓	✓		✓								✓								✓
Jackson & Mishra (2013)	✓‡‡	✓				✓				✓		✓	✓	✓	✓		✓	✓	✓	✓							✓	‡‡									
Janszky et al (2010)	✓§§	✓	✓		✓							✓	✓	✓	✓			✓	✓		✓				✓									✓			
Joyce (2015)	✓	✓	✓			✓ ***	✓ ***					✓	✓	✓	✓	✓	+✓	✓	✓		+✓						✓ ***		✓	✓						
Kim et al (2011)	✓	✓	✓	✓		✓												✓	✓		✓	✓														✓
Kim et al (2013)	✓	✓	✓	✓		✓		✓										✓			✓		✓													✓
Köhler et al (2013)	✓	✓	✓			✓						✓					✓	✓	✓		✓			✓												✓

Table 3.2 continued: Summary description of covariates in included studies (page 3 of 5)

	Sociodemographic factors											Lifestyle factors					Comorbidities/ impairment									Biological factors			Medication					Study factors		Familial factors		
Author	Age	Sex	Marital status	Race/ ethnicity	Region/ urbanisation	Education	Occupation/ employment status	Income/ wealth	Area-based deprivation	Homeownership	Social support/ social deprivation	Smoking	Physical inactivity	Alcohol use	BMI/obesity/weight/ waist circumference	Diet/ dietary supplements	CVD (other than stroke/MI)	Blood pressure/ hypertension	Diabetes (mellitus)	Cholesterol/ lipid levels	Cancer	Psychopathology (other than depression)	Physical impairment	Cognitive impairment/ function	Height	C-reactive protein	Menopausal status/ HRT use	Medication adherence	Antidepressants/ antipsychotics	Antihypertensive medication	Lipid-/ cholesterols-lowering drugs	Aspirin/ antithrombotic drugs	Study site	Year of enrolment	Mother's/ father's SES	Family history of CVD/ CHD	**Others***
Kubzansky et al (2006)	✓	✓				✓						✓		✓	✓		✓†	✓	✓†	✓		✓													✓	✓
Langvik & Hjemdal (2015)	✓	✓										✓			✓		✓†	✓	✓	✓		✓														
Larson et al (2001)	✓	✓		✓		✓	✓ ‡‡‡	✓ ‡‡‡				✓					✓	✓	✓			✓†														
Lin et al (2014)	✓	✓															✓	✓	✓	✓	✓															
Majed et al (2012)	✓	✓	✓			✓	✓					✓	✓	✓	✓		✓†	✓		✓									✓	✓			✓			✓
Marijnissen et al (2014)	✓	✓										✓			✓		✓‡	✓	✓				✓	✓												✓
Mathur et al (2016)	✓	✓		✓					✓			✓			✓			✓	✓	✓		✓							✓							
Mejía-Lancheros et al (2014)	✓	✓										✓		✓	✓		✓†	✓	✓	✓	✓†	✓†													✓	✓
Moise et al (2016)	✓	✓	✓		✓	✓		✓				✓	✓	✓	✓		✓†	✓	✓	✓						✓		✓	✓	✓	✓	✓				✓
Niles & O'Donovan (2018)	✓	✓	✓	✓		✓						✓	✓	✓	✓							✓														
O'Brien et al (2015)	✓ §§§	✓ §§§				✓		✓				✓ §§§	✓	✓	✓ §§§		✓ §§§	✓ §§§	✓ §§§										✓	✓ §§§		✓ §§§					✓
Ohira et al (2001)	✓	✓										✓		✓	✓		✓†	✓	✓	✓										✓						
Pan et al (2011a)	✓ ‡‡	✓	✓ ‡‡	✓								✓ ‡‡	✓ ‡‡	✓ ‡‡	✓ ‡‡	✓ ‡‡	✓ ‡‡	✓ ‡‡	✓ ‡‡	✓ ‡‡	✓ ‡‡						✓ ‡‡					✓ ‡‡				✓	

Table 3.2 continued: Summary description of covariates in included studies (page 4 of 5)

Author	Sociodemographic factors											Lifestyle factors					Comorbidities/ impairment									Biological factors			Medication					Study factors		Familial factors		Others*
	Age	Sex	Marital status	Race/ ethnicity	Region/ urbanisation	Education	Occupation/ employment status	Income/ wealth	Area-based deprivation	Homeownership	Social support/ social deprivation	Smoking	Physical inactivity	Alcohol use	BMI/obesity/weight/ waist circumference	Diet/ dietary supplements	CVD (other than stroke/ MI)	Blood pressure/ hypertension	Diabetes (mellitus)	Cholesterol/ lipid levels	Cancer	Psychopathology (other than depression)	Physical impairment	Cognitive impairment/ function	Height	C-reactive protein	Menopausal status/ HRT use	Medication adherence	Antidepressants/ antipsychotics	Antihypertensive medication	Lipid-/ cholesterols-lowering drugs	Aspirin/ antithrombotic drugs	Study site	Year of enrolment	Mother's/ father's SES	Family history of CVD/ CHD	
Péquignot et al (2013)	✓	✓				✓					✓	✓		✓			+✓	✓	✓	✓				(+)	✓								✓				
Péquignot et al (2016)	✓	✓	✓			✓					✓	✓		✓			+✓	✓	✓	✓				(+)	✓								✓				
Pössel et al (2015)	✓	✓										✓					+✓	✓		✓																	✓
Pratt et al (1996)	✓	✓	✓			✓ ****		✓ ****				✓ ****		✓			+✓	✓	✓ ****			✓ ****						✓									
Scherrer et al (2010)	✓																+✓					+✓											✓				
Scherrer et al (2011)	✓	✓	✓	✓													+✓					+✓											✓				
Sesso et al (1998)	✓	✓	✓									✓		✓	✓		+✓	✓	+✓															✓				
Stewart et al (2016)	✓	✓		✓								✓			✓		+✓	✓	✓	✓		✓															
Stürmer et al (2006)	✓	✓				✓						✓	✓	✓	✓		✓	✓	✓	✓	✓														✓			
Sun et al (2016)	✓	✓	✓		✓	✓		✓				✓	✓	✓	✓		+✓	✓	✓	✓	+✓	(✓)														✓		
Wassertheil-Smoller et al (2004)	✓	✓		✓		✓		✓				✓	✓		✓		+✓	✓		✓		+✓		✓			✓										
Wouts et al (2008)	✓	✓				✓						✓					‡✓	§✓					§✓	✓			(+)										

Table 3.2 continued: Summary description of covariates in included studies (page 5 of 5)

	Sociodemographic factors											Lifestyle factors					Comorbidities/ impairment								Biological factors			Medication					Study factors		Familial factors		
Author	Age	Sex	Marital status	Race/ ethnicity	Region/ urbanisation	Education	Occupation/ employment status	Income/ wealth	Area-based deprivation	Homeownership	Social support/ social deprivation	Smoking	Physical inactivity	Alcohol use	BMI/obesity/ weight/ waist circumference	Diet/ dietary supplements	CVD (other than stroke/ MI)	Blood pressure/ hypertension	Diabetes (mellitus)	Cholesterol/ lipid levels	Cancer	Psychopathology (other than depression)	Physical impairment	Cognitive impairment/ function	Height	C-reactive protein	Menopausal status/ HRT use	Medication adherence	Antidepressants/ antipsychotics	Antihypertensive medication	Lipid-/ cholesterols-lowering drugs	Aspirin/ antithrombotic drugs	Study site	Year of enrolment	Mother's/ father's SES	Family history of CVD/ CHD	**Others***
Yan et al (2013)	✓	✓	✓			✓		✓																												

* More details about additional covariates considered in included studies are provided in Appendix Table A 8
† Restriction
‡ Stratification
§ In some analyses included as time-dependent variable
** Variable was taken into account through direct inclusion into MSM model (baseline age) and by applying inverse probability weights (age at interview)
†† Variable was taken into account by applying inverse probability weights
‡‡ Time-dependent covariate
§§ Through study design (participants almost the same age: 18 – 20 years at baseline)
*** Added in stepwise but did not end up in final model
††† Added as index of chronic illness (self-reported doctor diagnosis of high blood pressure, cancer, lung disease, psychiatric problems, and arthritis)
‡‡‡ Nam-Powers Index (SES using income, occupation, prestige ratings)
§§§ These variables were taken into account by fitting a separate model to obtain a stroke risk score which in turn was included in the regression model
**** Variable was removed from the model since it was not significant at $p < 0.10$ and removal did not alter OR by >20%
CVD: cardiovascular diseases, CHD: coronary heart disease, HRT: hormone replacement therapy, MI: myocardial infarction; (): only taken into account in some of the analyses

3.4.3 Quality assessment

The assessment of the included studies' focus, recruitment strategy, accuracy of exposure and outcome, method of conditioning on confounding and mediating factors, length and loss of follow-up, precision and generalisability of results allowed for an assessment of the strengths and weaknesses of included studies (see section 3.5.2 Strengths and limitations of included studies for detailed assessment). Overall, the included studies were rated to have a clear focus but there were important shortcomings in the exposure assessments, precision of results, and method of conditioning on confounding and mediating factors. An example of a completed CASP Cohort Study Checklist is included in Appendix 6.

3.4.4 Data synthesis

3.4.4.1 Stroke

3.4.4.1.1 Assessing eligibility of studies for inclusion in meta-analysis

Among 36 studies on the association between depression and risk of stroke, 32 studies explored the risk of total stroke. Twenty-four of these studies were eligible for inclusion in the meta-analysis. Reasons for exclusion were overlapping populations (Gilsanz et al, 2015; Glymour et al, 2012; Marijnissen et al, 2014; Péquignot et al, 2013) and non-comparable effect estimates (Jackson & Mishra, 2013; Kim et al, 2011; Kim et al, 2013; Larson et al, 2001). Twenty-three eligible studies reported effect estimates on the hazard of total stroke among individuals with depression or depressive symptoms relative to non-exposed individuals. Additionally, seven studies reported the hazard of total stroke per unit increase on a depressive symptom rating scale. Ten studies investigated the hazard of ischaemic stroke of which eight studies were eligible for inclusion whereas two studies were excluded due to overlapping populations (Gilsanz et al, 2017; Yan et al, 2013). Out of six studies that examined the hazard of haemorrhagic stroke five studies were eligible for inclusion in the meta-analysis. One study was excluded because authors did not provide CIs (Henderson et al, 2013). Four of the included studies reported the hazard of all types of haemorrhagic stroke whereas Daskalopoulou et al (2016) reported results separately for subarachnoid and intracerebral haemorrhage. Since both effect estimates were

based on the same individuals and the number of people who had an intracerebral haemorrhage was higher, the decision was made to only include the effect estimate on the hazard of intracerebral haemorrhage in the meta-analysis. The estimate of the association between depression and hazard of unknown stroke type reported by Pan et al (2011a) was extracted but was not included in any of the meta-analyses.

3.4.4.1.2 Total stroke - dichotomous exposure

Twenty-three studies reported effect estimates of the hazard of total stroke among individuals with depression or depressive symptoms relative to non-exposed individuals (Appendix Table A 2). The reported HRs ranged from 0.44 (95% CI: 0.06 – 3.22) to 8.50 (95% CI: 1.05 – 68.73) with these extreme values based on sub-groups of people in each study (Gafarov et al, 2017; Wouts et al, 2008). Nine studies reported effect estimates that did not overlap with the null value, all of which indicated increased hazards of total stroke among individuals with depression or depressive symptoms relative to non-exposed individuals. The pooled random effects meta-analysis indicated a 30% (95% CI: 19 – 43%) increased hazard of stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals (Figure 3.2). There were statistically significant differences between studies and the estimate of between-study heterogeneity that remained unexplained was moderate (I^2: 38.2%).

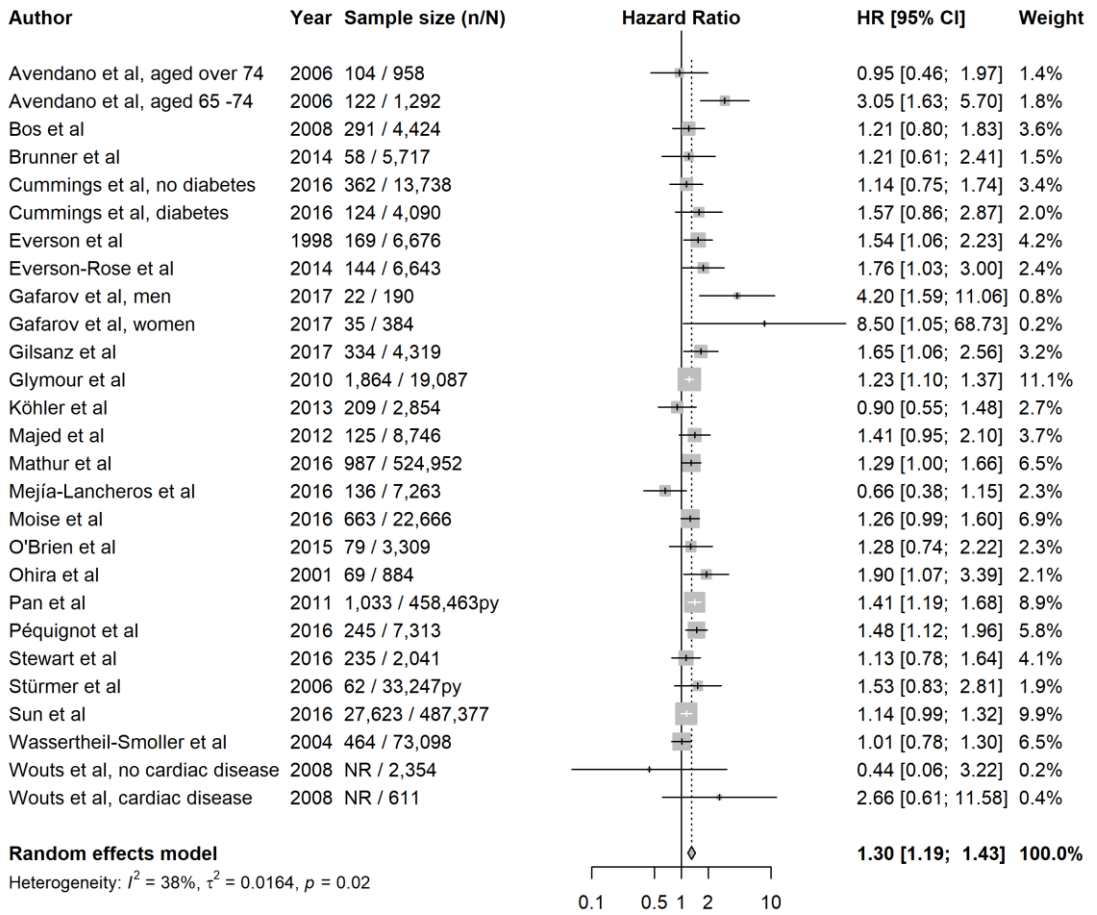


Figure 3.2: Meta-analysis of the hazard of total stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals

Among studies that provided sex-specific estimates, depression or depressive symptoms were associated with increased hazards of stroke among both men and women (pooled HR: 1.26, 95% CI: 1.01 – 1.57, and 1.23, 95% CI: 1.06 – 1.44, respectively) (Figure 3.3). There was moderate between-study heterogeneity that remained unexplained among studies with male and female participants (I^2: 45.1% and 53.5%, respectively).

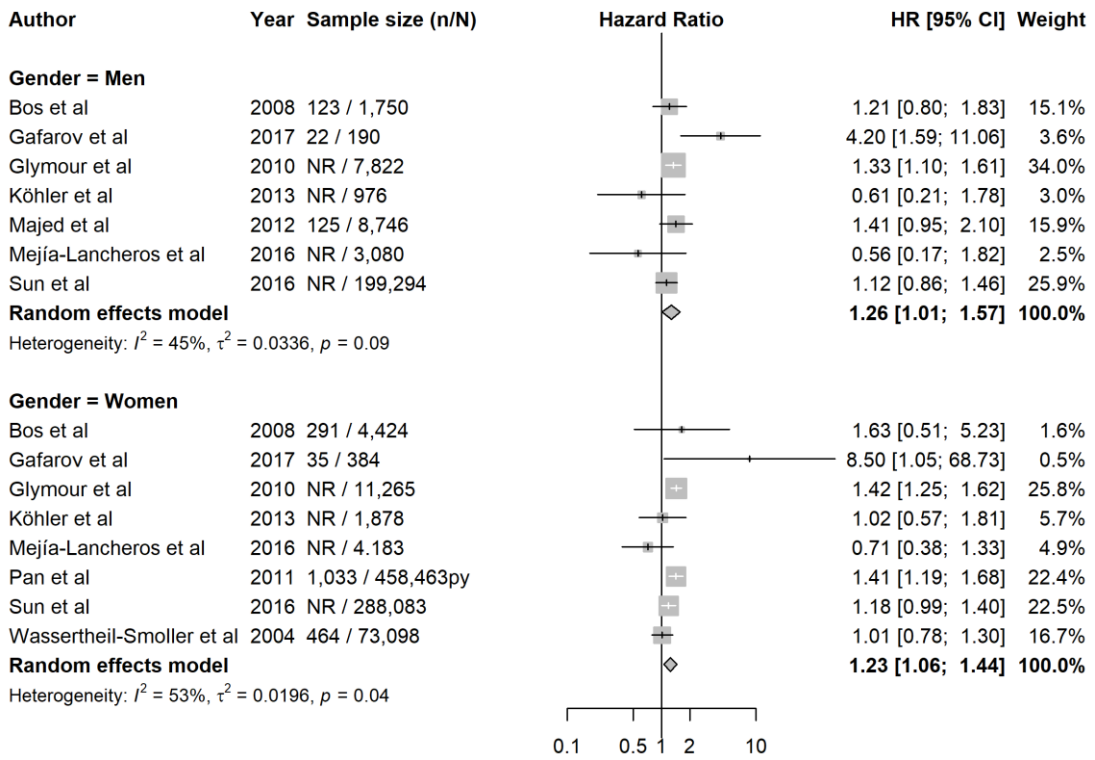


Figure 3.3: Meta-analysis of the hazard of total stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals, separately for men and women

The pooled effect estimates were slightly higher among studies using a measure of depressive symptoms than among studies using a measure of clinical depression (Figure 3.4). The pooled HR indicated a possible 20% increased hazard of stroke among individuals with clinical depression, relative to individuals without clinical depression (pooled HR: 1.20, 95% CI: 0.99 – 1.45). There were no statistically significant differences among studies using a measure of clinical depression (p = 0.06). However, the p-value of the chi^2 test of differences between studies was close to the cut-off value of statistical significance. The estimated proportion of heterogeneity that remained unexplained was moderate at 52.5%. The increased hazard of stroke among individuals with depressive symptoms was estimated to be 35% (95% CI: 21 – 50%), relative to individuals without depressive symptoms. Heterogeneity among studies using a measure of depressive symptoms was moderate (35.4%, p = 0.06).

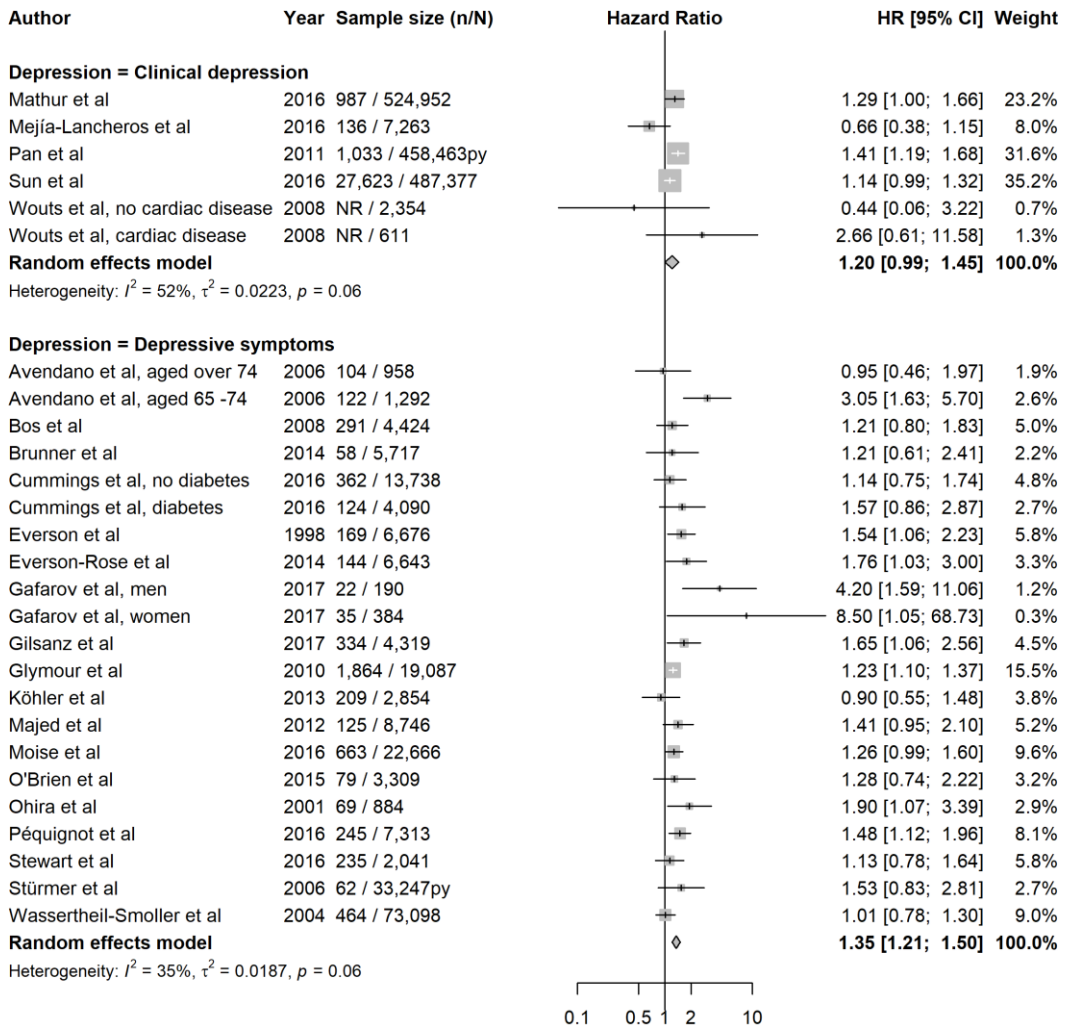


Figure 3.4: Meta-analysis of the hazard of total stroke among individuals with clinical depression or depressive symptoms, relative to non-exposed individuals, separately for studies using a measure of clinical depression and depressive symptoms

3.4.4.1.3 Total stroke - continuous exposure

Six studies investigated the hazard of total stroke assessing depressive symptoms on a continuous scale (Appendix Table A 3). Three studies investigated the hazard of stroke per one point increase on depressive symptom scales (Everson et al, 1998; Glymour et al, 2010; Wouts et al, 2008) and three studies reported the increased hazard per one standard deviation increase (O'Brien et al, 2015; Ohira et al, 2001; Stürmer et al, 2006). Ohira et al (2001) provided sufficient information to convert the effect size from per standard deviation to per unit increase whereas O'Brien et al (2015) and Stürmer et al (2006) did not provide sufficient information. The pooled HR per one point increase on depressive symptom scales was estimated at 1.05 (95% CI:

1.01 – 1.08) (Figure 3.5). Substantial between-study heterogeneity remained unexplained (I^2: 83.4%) which was unlikely to be due to chance alone ($p < 0.01$). The pooled HR per standard deviation increase on depressive symptom scales was estimated as 1.22 (95% CI: 1.02 – 1.45) (Figure 3.5). There was no between-study heterogeneity that remained unexplained (0.0%).

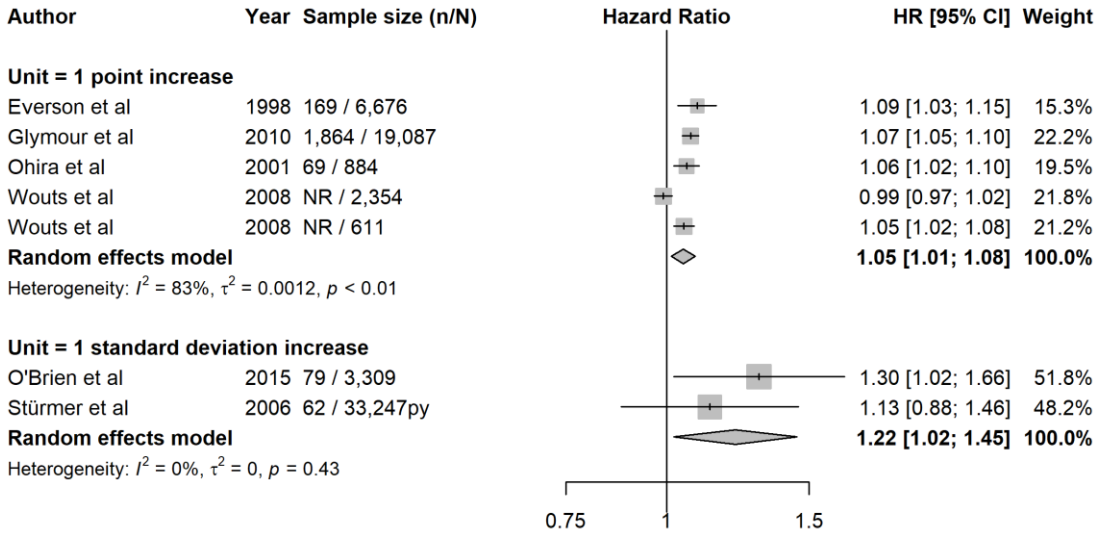


Figure 3.5: Meta-analysis of the hazard of total stroke per unit increase on depressive symptom scales

3.4.4.1.4 Ischaemic stroke

All eight included studies reported increased hazards of ischaemic stroke among exposed individuals relative to unexposed individuals with effect estimates ranging from 1.13 (95% CI: 0.96 – 1.34) to 2.70 (95% CI: 1.21 – 6.04) (Ohira et al, 2001; Sun et al, 2016) (Appendix Table A 4). The pooled HR was estimated at 1.27 (95% CI: 1.13 – 1.34) (Figure 3.6). There were no statistically significant differences between studies (p = 0.23) and between-study heterogeneity was relatively low (I^2: 24.8%).

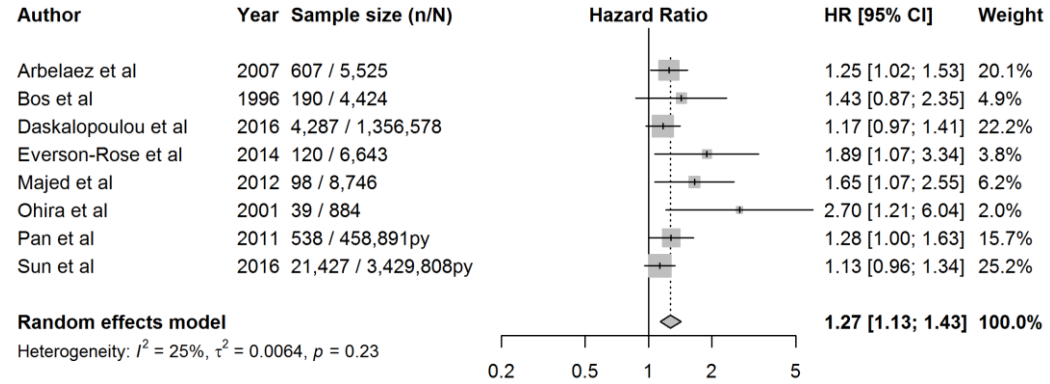


Figure 3.6: Meta-analysis of the hazard of ischaemic stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals, separately for studies using a measure of clinical depression and depressive symptoms

3.4.4.1.5 Haemorrhagic stroke

Whilst four of the five HRs indicated a possible increased hazard of haemorrhagic stroke among individuals with the exposure relative to individuals without depression or depressive symptoms, the estimates of included studies were imprecise and all reported CIs included the null value (Appendix Table A 5). The pooled HR was estimated at 1.17 (95% CI: 0.97 – 1.42) (Figure 3.7). There were no statistically significant differences between studies and no between-study heterogeneity remained unexplained (I^2: 0.0%).

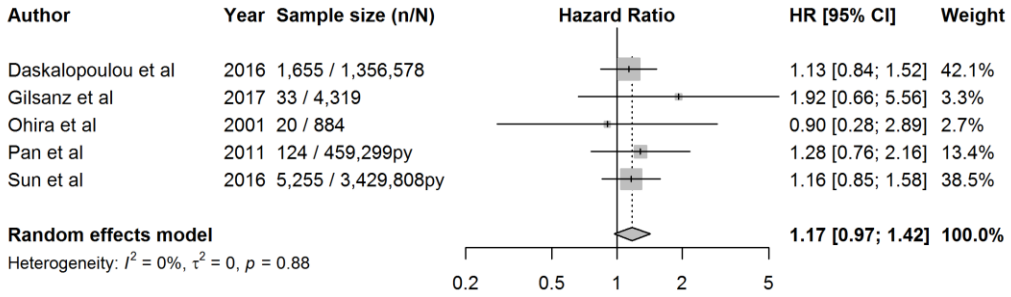


Figure 3.7: Meta-analysis of the hazard of haemorrhagic stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals, separately for studies using a measure of clinical depression and depressive symptoms

3.4.4.1.6 Summary of main findings of studies not included in the meta-analysis

3.4.4.1.6.1 Total stroke

The results of excluded studies generally followed the same pattern as the overall pooled estimate from included studies, with the majority of studies indicating

increased risks of individuals with depression or depressive symptoms, relative to unexposed individuals.

Eight excluded studies reported estimates on the risk of total stroke among individuals with depression or depressive symptoms relative to individuals without depression or depressive symptoms (Figure 3.8). All studies that were excluded due to the provision of non-comparable effect estimates reported increased risks of total stroke among exposed individuals relative to unexposed individuals (Jackson & Mishra, 2013; Kim et al, 2011; Kim et al, 2013; Larson et al, 2001). Whilst the magnitude of the effects was not directly comparable between excluded studies and the meta-analysis, the direction of the effects were the same. Three of the four studies that were excluded due to overlapping populations showed increased hazards of total stroke among exposed individuals relative to those without depression or depressive symptoms (Gilsanz et al, 2015; Marijnissen et al, 2014; Péquignot et al, 2013). The CIs of the estimates by Gilsanz et al (2015) and Marijnissen et al (2014) did not overlap with the CI of the pooled HR indicating that the observed effect estimates were more different than would be expected due chance alone. One potential reason for the differences between the pooled effect estimate and the estimate of the study by Gilsanz et al (2015) was the difference in the exposure assessment. The reported effect estimate reflected the hazard of total stroke among individuals with stable high depressive symptoms relative to those with stable low depressive symptoms. The effect estimates were lower among individuals with recent onset depressive symptoms (1.08, 95% CI: 0.81 – 1.44) and recently remitted depressive symptoms (1.66, 95% CI: 1.22 – 2.26), relative to those with stable low depressive symptoms. However, these estimates were not included in the meta-analysis due to overlap in participants in the comparison groups. Another potential explanation is methodological differences in dealing with time-updated covariates and loss to follow-up. The estimate provided by Marijnissen et al (2014) was imprecise and the number of events was low indicating that the results might be influenced by random error. One study showed an increased hazard of total stroke among White adults but no association among Hispanic and African-American adults (Glymour et al, 2012).

However, the estimates of the latter two groups were imprecise and CIs overlapped with the null value.

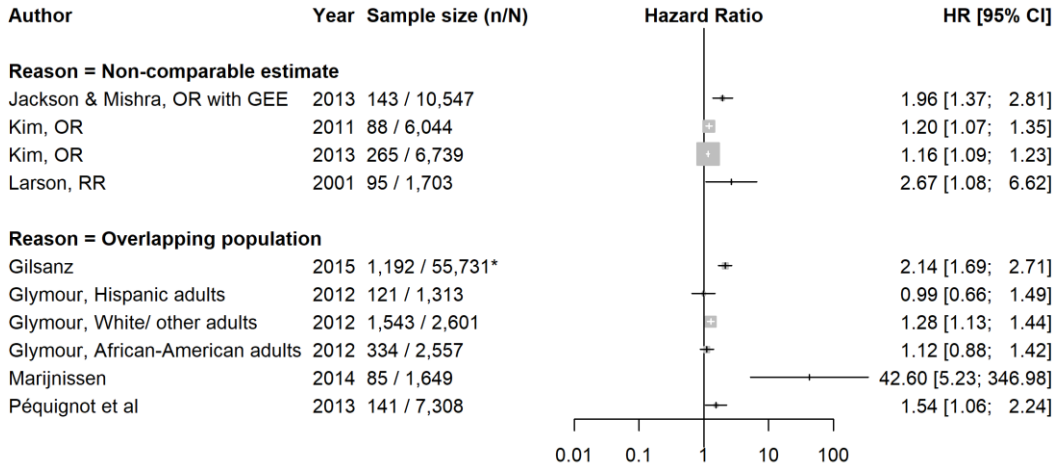


Figure 3.8: Results of studies not included in the meta-analyses on the risk of stroke

The direction of the effect of both excluded studies on the risk of total stroke per unit increase on a depressive symptom rating scale is in keeping with studies included in the meta-analysis (Marijnissen et al, 2014; Niles & O'Donovan, 2018). Marijnissen et al (2014) reported a 12% (95% CI: 3 – 22%) increased hazard of fatal or non-fatal stroke per unit increase on the depressive symptom scale. Similarly, Niles & O'Donovan (2018) found higher odds of fatal stroke per unit increase on the depressive symptom scale (OR: 1.12, 95% CI: 1.01 – 1.24). However, the authors did not report what unit of increase was used.

3.4.4.1.6.2 Ischaemic stroke

The studies by Gilsanz et al (2017) and Yan et al (2013) were excluded from the meta-analysis on the risk of ischaemic stroke because they were based on the same data source as the study by Arbelaez et al (2007). Whilst Gilsanz et al (2017) reported higher estimates than Arbelaez et al (2007) (1.64, 95% CI: 1.04 – 2.60 and 1.25, 95% CI: 1.02 – 1.53, respectively), Yan et al (2013) reported slightly lower effect estimates for both White and African American adults (1.20, 95% CI: 0.97 – 1.48, and 1.21, 95% CI: 0.97 – 1.84, respectively). However, the CIs of all estimates overlapped. As discussed previously, one potential explanation is the difference between exposure assessments

and the methods of conditioning on covariates. Of note, the highest risk of ischaemic stroke was observed among individuals with stable high depressive symptoms, relative to those with stable low depressive symptoms. The corresponding estimates for recent onset depressive symptoms and recently remitted depressive symptoms were 1.49 (95% CI: 0.97 – 2.30) and 1.08 (95% CI: 0.69 – 1.70), respectively (Gilsanz et al, 2017).

3.4.4.1.6.3 Haemorrhagic stroke

The study by Henderson et al (2013) was excluded from the meta-analysis on the risk of haemorrhagic stroke since authors did not provide CIs. The reported HR was higher than the effect estimates of any included studies (HR: 1.63, $p < 0.01$). It is likely that one reason for the high hazard ratio is that a different unit of measurement was used. Whilst all included studies dichotomised the exposure, Henderson et al (2013) assessed the risk of haemorrhagic stroke per standard deviation increase on the depressive symptom scale.

3.4.4.2 Myocardial infarction

3.4.4.2.1 Assessing eligibility of studies for inclusion in meta-analysis

Nineteen of the 24 studies that estimated the risk of MI were eligible for inclusion in the meta-analysis. The reasons for exclusion were non-comparable effect estimates (Kubzansky et al, 2006; Langvik & Hjemdal, 2015; Pratt et al, 1996) and overlapping populations (Lin et al, 2014; Scherrer et al, 2011). Sixteen of the eligible studies reported the hazard of MI among individuals with depression or depressive symptoms relative to individuals without depression or depressive symptoms. Four studies estimated the hazard of MI using continuous depressive symptoms scales.

3.4.4.2.2 Dichotomous exposure

The reported hazards of MI among individuals with depression or depressive symptoms relative to those not exposed ranged from 0.80 (95% CI: 0.44 – 1.44) to 3.47 (95% CI: 0.51 – 23.64) (Gafarov et al, 2017; Stürmer et al, 2006) (Appendix Table A 6). Four of the studies reported CIs that did not overlap with the null value. Of these, all studies reported increased hazards of MI among exposed individuals relative to non-

exposed individuals. The pooled random effects HR of developing MI among individuals with depression or depressive symptoms to non-exposed individuals was 1.19 (95% CI: 1.08 – 1.31) (Figure 3.9). The chi² test showed statistically significant differences between studies that were unlikely due to chance alone ($p < 0.01$). Furthermore, the I^2 statistic suggested that a moderate to substantial part of the heterogeneity between studies remained unexplained ($I^2 = 65.9\%$).

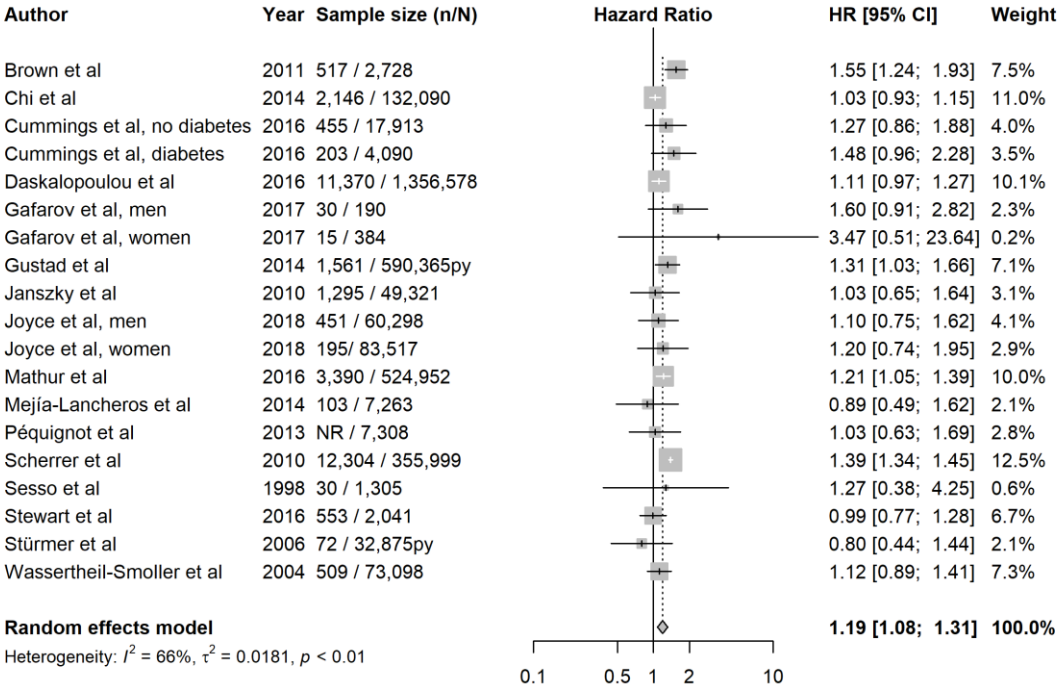


Figure 3.9: Meta-analysis of the hazard of myocardial infarction among individuals with depression or depressive symptoms, relative to non-exposed individuals

Depression or depressive symptoms were associated with increased hazards of MI among men and women (pooled HR, 95% CI: 1.12, 1.01 – 1.24, and 1.14, 1.04 – 1.26, respectively) (Figure 3.10). Both I^2 statistics indicated low heterogeneity between seven studies among men and six studies among women (I^2: 0.0%, and I^2: 0.0%, respectively). Furthermore, the chi² test suggested that differences between studies might be due to chance alone (men: $p = 0.68$, women: $p = 0.91$). This might indicate that sex is a potential source of heterogeneity. However, it should be noted that some studies were excluded from this subgroup analysis because they did not provide results stratified by sex.

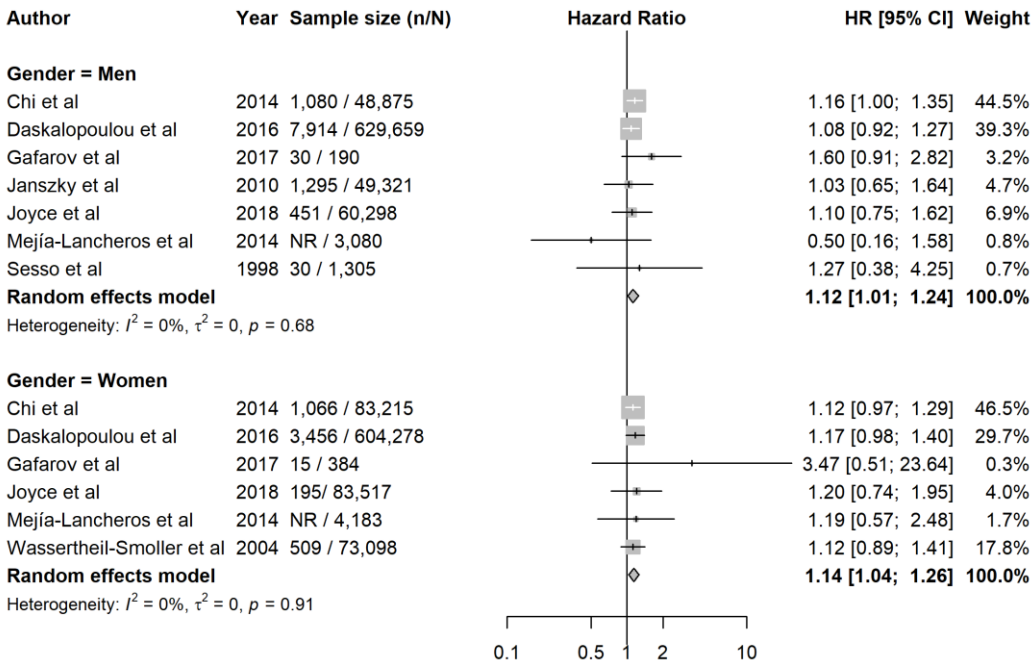


Figure 3.10: Meta-analysis of the hazard of myocardial infarction among individuals with depression or depressive symptoms, relative to non-exposed individuals, separately for men and women

There were increased hazards among individuals with clinical depression (pooled HR: 1.15, 95% CI: 1.00 – 1.32) as well as among individuals with depressive symptoms (pooled HR: 1.24, 95% CI: 1.09 – 1.41) relative to non-exposed individuals (Figure 3.11). Chi² tests indicated that statistically significant differences between studies remained among studies using a measure of clinical depression ($p < 0.01$), whereas differences between studies using a measure of depressive symptoms might be due to chance alone ($p = 0.22$). The proportion of heterogeneity that remained unexplained was substantial among studies using a clinical depression measure (I^2: 82.0%) but low among studies using a depressive symptom measure (I^2: 23.6%).

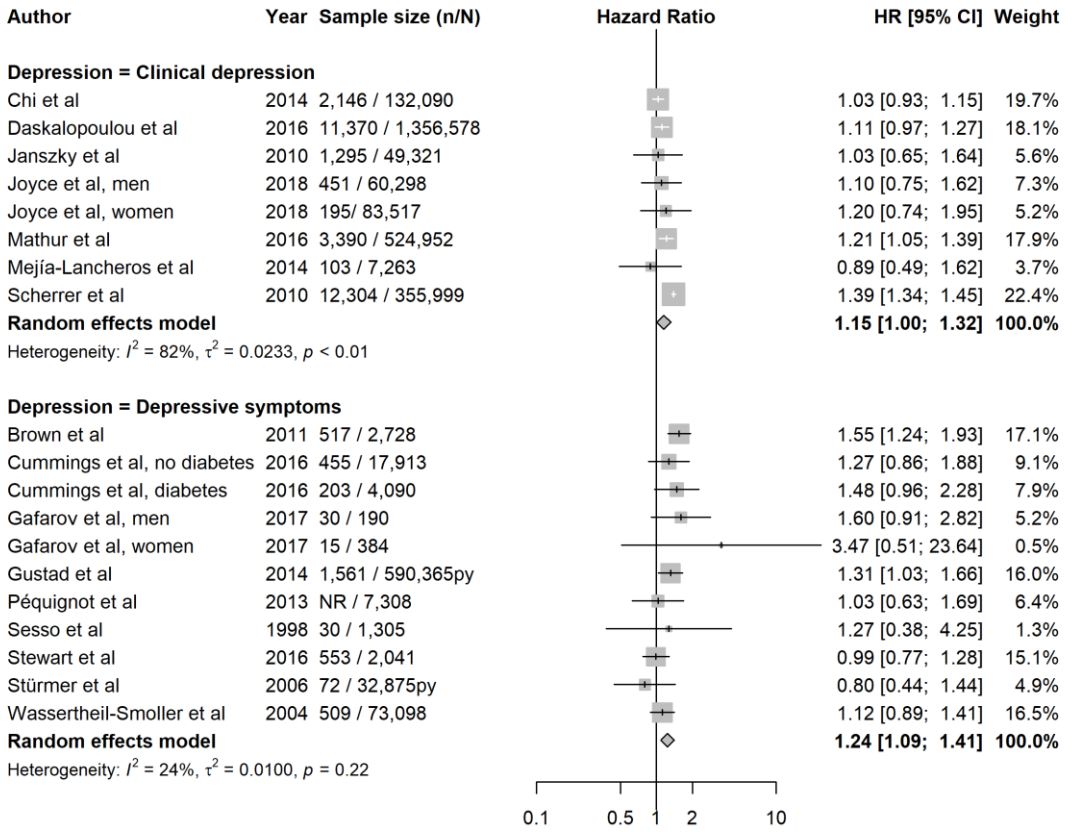


Figure 3.11: Meta-analysis of the hazard of myocardial infarction among individuals with depression or depressive symptoms, relative to non-exposed individuals, separately for studies using a measure of clinical depression and depressive symptoms

3.4.4.2.3 Continuous exposure

Four studies reported the hazard of MI per unit increase on depressive symptom scales (Appendix Table A 7). Whilst Pössel et al (2015) investigated the hazard of MI per one score increase on the depressive symptom scale, Ariyo et al (2000) investigated the hazard of MI per five unit increase on the depressive symptom scale, and Stürmer et al (2006) and Barefoot & Schroll (1996) investigated the hazard per one and two standard deviation increase on the depressive symptom scale, respectively. Ariyo et al (2000) and Barefoot & Schroll (1996) provided sufficient information to convert the reported effect size to the effect per one point increase on depressive symptom scales. Stürmer et al (2006) did not provide sufficient information. The pooled HR of MI per one point increase on depressive symptom scales was estimated at 1.04 (95% CI: 1.02 – 1.06) (Figure 3.12). There was no heterogeneity between studies

(I^2: 0.0%, p = 0.45). Stürmer et al (2006) reported an increased hazard of 1.09 (95% CI: 0.84 – 1.41) per standard deviation increase on the depressive symptom scale.

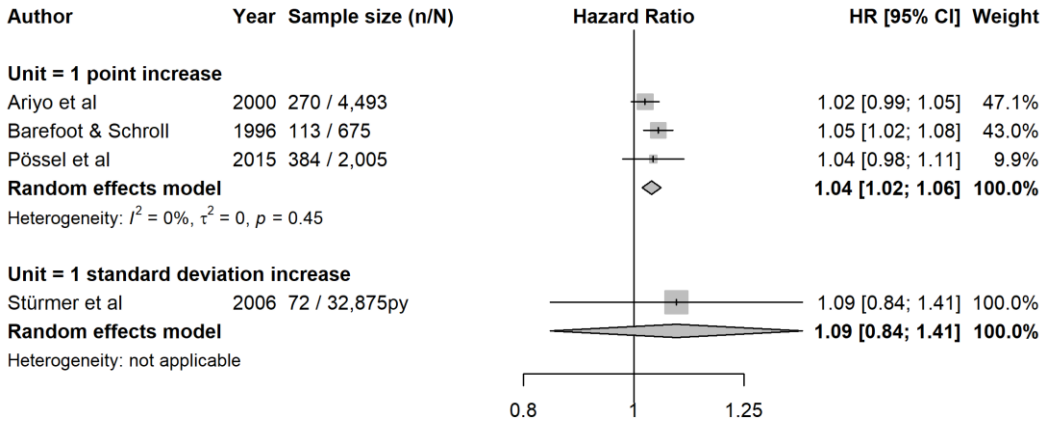


Figure 3.12: Meta-analysis of the hazard of myocardial infarction per unit increase on depressive symptom scales

3.4.4.2.4 Summary of main findings of studies not included in the meta-analysis

Overall, the results of excluded studies followed the same pattern as results of included studies, but the majority of excluded studies were small with low precision of estimates. Four of the five excluded studies reported effect estimates of the risk of MI among individuals with depression or depressive symptoms relative to non-exposed individuals (Figure 3.13). The point estimates of all excluded studies were higher than the point estimate of the pooled HR. Two studies were excluded due to non-comparable effect estimates. Pratt et al (1996) reported increased odds of MI among individuals with depressive symptoms relative to individuals without depressive symptoms. However, the number of events in the exposure and comparison groups were low (n = 6 and 37, respectively) and, as a result, the estimate was very imprecise. Kubzansky et al (2005) reported an increased hazard of MI among individuals with depression in the highest tertile relative to individuals in the middle tertile but the estimate was imprecise and CIs overlapped with the null value. Two studies were excluded because they were based on the same data source as one included study. Both studies reported increased hazards of MI among individuals with depressive symptoms relative to individuals without depressive symptoms which is in keeping with the pooled HR; however, in both studies the magnitude of

the effect was higher. Whilst CIs of the study by Scherrer et al (2011) overlapped with those of the pooled HR, there was no overlap of the CIs among the study by Lin et al (2014). The study that was based on the same data source used the same exposure and outcome assessment but reported a lower effect estimate (1.03, 95% CI: 0.93 – 1.15) (Chi et al, 2014). One potential reason might be differences between the study populations. However, the age and sex distribution and selection procedure of the two studies were very similar. Another potential reason is differences in length of follow-up. Whilst Lin et al (2014) followed participants for ten years, the length of follow-up in the study by Chi et al (2014) was six years.

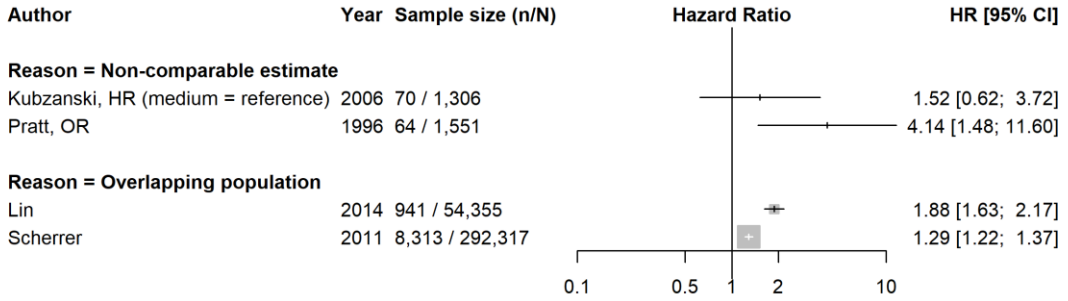


Figure 3.13: Results of studies not included in the meta-analyses on the risk of myocardial infarction

One of the excluded studies estimated the odds of non-fatal MI per unit increase on the depressive symptom scale (OR: 1.04, 95% CI: 1.01 – 1.04) (Langvik & Hjemdal, 2015). Whilst the direction of the effect is the same as the effect observed in the meta-analysis, the magnitude of the effect is not directly comparable to the pooled effect estimate of the meta-analysis due to the use of odds ratios (ORs) instead of HRs.

3.5 Discussion

3.5.1 Summary of key findings

The meta-analysis of 51 studies suggested that the hazard of stroke and MI was higher among individuals with depression or depressive symptoms relative to non-exposed individuals. The association was slightly stronger for stroke than for MI, although CIs overlapped. For both outcomes, the results were similar for men and women. Furthermore, there were increased hazards of total stroke, ischaemic stroke, and haemorrhagic stroke among individuals with depression or depressive symptoms

relative to non-exposed individuals. However, the effect estimate for the association between depression and haemorrhagic stroke was imprecise and overlapped with the null value. There might be a dose-dependent effect of depressive symptoms on hazard of MCVE since there were increased hazards of stroke and MI per unit increase on depressive symptom rating scales. In contrast, the point estimates for hazards of stroke and MI were higher among studies using a measure of depressive symptoms than among studies using a measure of clinical depression. Again, CIs largely overlapped. The results of studies excluded from the meta-analysis were consistent with results of included studies in that the majority of studies showed increased risks of MCVE among individuals with depression or depressive symptoms. There were some differences in the effect sizes which might have been explained by differences between samples, methodology, or the effect measure used. Since the results of meta-analyses are directly influenced by the strengths and weaknesses of included studies, the following will discuss to what extent chance, bias, and confounding might explain the increased risk of stroke or MI among individuals with depression and depressive symptoms. After that, strengths and weaknesses of this systematic review and meta-analysis are presented, the findings are put in the context of existing literature, and implications for research are given. Implications for practice are presented at the end of this thesis, taking into account the results of all projects (see section 7.5 Implications for practice).

3.5.2 Strengths and limitations of included studies

3.5.2.1 Chance

Some included studies are very likely influenced by chance. All but one of the studies considered the possibility of chance by providing CIs. Wide CIs indicated imprecise estimates in a number of included studies. This could either be explained by small sample sizes or by small numbers of events in either the exposure or comparison groups. For example, even though the overall sample size (N) and number of events (n) of the study by Janszky et al (2010) was fairly largely (n/ N: 1,295 / 49,321), the number of events and the size of the group with clinical depression were small (n/ N: 22 / 646) which resulted in imprecise estimates. As a result, it cannot be established

with certainty whether there truly was no association between depression and MI or whether the findings were explained by limited power. Also, some studies provided estimates separately for subgroups. For example, Wouts et al (2008) provided results separately for individuals with and without cardiac diseases and Cummings et al (2016) provided results separately for individuals with and without diabetes. Whilst the point estimates between these subgroups differed, the CIs of the estimates largely overlapped indicating that differences between groups might be due to chance alone. Henderson et al (2013) provided a p-value as measure of statistical significance but did not provide a CI. Whilst, the authors showed that there is a statistically significant association between depression and haemorrhagic stroke, it was not possible to assess the precision of the estimate.

3.5.2.2 Bias

3.5.2.1 Selection bias

The selection of participants might have biased the results of included studies. First, results might be biased if the sample did not represent the population of interest. The population of interest were adults free from stroke and/ or MI at baseline. The exclusion of participants with prior stroke or MI was important since a history of stroke or MI would raise questions of reverse causation in that a prior history of stroke or MI might increase the risk of depression as well as the risk of future cardiovascular events. Included studies aimed to exclude adults with a history of stroke or MI relying on self-report and/ or medical records. Some studies might have missed participants with a history of stroke and/ or MI, particularly if they relied on self-report. Also, studies differed in their extent of excluding participants with prior CVD other than stroke or MI. Therefore, the results of some studies might partly reflect reverse causation.

Second, results might be biased if the exposure group differs from the comparison group in factors other than the exposure of interest. All studies sampled both groups from the same source population. However, twenty-three studies reported baseline differences between the exposure and comparison group in characteristics other than

the exposure of interest. These differences were considered in analyses of studies through restriction, stratification, adjustment and inverse probability weighting. The impact of applying these different methods will be further discussed in section 3.5.2.3 Confounding and mediation.

Third, selection bias might have influenced findings due to differential loss to follow-up among the exposure and comparison groups. Since information of participants lost to follow-up was unknown, studies inevitably conditioned on loss to follow-up by restricting their analysis to those not lost to follow-up. As a result, the effect estimates represented the association between depression or depressive symptoms and risk of stroke or MI among those not lost to follow-up whilst effect estimates among those lost to follow-up remained unknown. Since individuals who were lost to follow-up often differ from those not lost to follow-up, this might introduce selection bias. However, whilst loss to follow-up is often a concern in prospective studies, it was not a major concern in the included studies, at least for those conducted among stable populations, because linkage to hospital and death records was used to follow participants over time.

Fourth, competing risks during follow-up might have introduced selection bias since the occurrence of the competing event (death from causes other than stroke or MI) made it impossible for the event of interest to occur and it was likely related to the exposure of interest. Therefore, it is important to take competing risks into account at the analysis stage. However, the potential impact of competing events were taken into account in three studies only (Gilsanz et al, 2017; Gilsanz et al, 2015; Köhler et al, 2013). Two of these studies applied stabilised inverse probability weights because of which the exposure was no longer associated with the competing risk, therefore preventing an open backdoor path that might bias the results (Gilsanz et al, 2017; Gilsanz et al, 2015). One study calculated a subdistribution hazard ratio (SHR) on the association between depression and stroke whilst considering death as competing event (Köhler et al, 2013). This analysis indicated a possible 20% decreased hazard of stroke among individuals with depressive symptoms in fully adjusted models (SHR:

0.80, 95% CI: 0.47 – 1.36). However, it should be noted that the study by Köhler et al (2013) was one of only a few studies that did not observe an increased hazard of stroke among individuals with depressive symptoms in analyses not considering competing risks.

3.5.2.2 Information bias

3.5.2.2.1 Exposure assessment

Inaccurate measurement of the exposure has very likely introduced information bias. First, some studies assessed depressive symptoms based on rating scales whilst others used measures of clinical depression based on medical records or self-report. Whilst self-reported information was prone to recall and response bias, studies using information from medical records were only able to define participants as depressed if participants had access to and used health care services. Considering that a large proportion of studies were conducted in the US where health services are not free of charge, this might have introduced bias. If depression is indeed associated with increased risk of cardiovascular events, this would have biased the estimate towards null. This bias might also explain why the pooled effect estimates of studies using a measure of clinical depression were lower than those of studies using a measure of depressive symptoms. However, there were other potential explanations of this finding. For example, the comparison groups of studies using a measure of clinical depression included participants with depressive symptoms. If depressive symptoms indeed increased the risk of stroke and MI, this would have biased the effect measure of these studies towards null. Another potential explanation is that observed differences were due to chance alone since the CIs of both estimates largely overlapped. It seems likely that a non-causal explanation is responsible for the higher effect estimate among studies using a measure of depressive symptoms instead of clinical depression since it appears counterintuitive that less exposure is associated with more harm.

Second, most but not all of the studies used rating scales that have been validated and some studies have used validated scales but have used different cut-points to allocate

participants to the exposure and comparison groups. Examples of well-validated scales are the CES-D (Orme et al, 1986; Radloff, 1977) and the Hospital and Anxiety Depression Scale (HADS) (Bjelland et al, 2002; Snaith, 2002). In contrast, to the best of my knowledge, the HPL Depression scale and the MONICA psychosocial depression scale have not been validated. The use of scales that have not been validated is problematic since it remained unknown how well these scales were able to differentiate between individuals with high and low measures of depressive symptoms. The use of different cut-points of the same scale hampered the comparison of findings across studies since it remained unknown to what extent differences were due to differences in chosen cut-points, sample variation, or true differences in findings. Also, some studies used cut-points that were data-driven. For example, Everson-Rose et al (2014) created groups based on quartiles of the CES-D score, the top group was further split into two groups, and the group with the highest scores was compared to the group with the lowest scores. Since the cut-points were data-driven, the distribution of depressive symptoms in the sample greatly influenced to what extent the exposure and comparison group differed in the exposure of interest. For example, if there was great variation in depressive symptom scores, then this method would compare two groups with very high and low depressive symptom scores. In contrast, if there was less variation in depressive symptom scores, then the exposure of interest would be similar among the two groups. Furthermore, the use of data-driven cut-points reduced the generalisability of results. Moreover, some studies used continuous scales. Whilst this was beneficial in terms of increased statistical power compared to categorical definitions of exposure, it might have been problematic if the scale was not validated to assess symptom severity.

Third, most studies allocated participants to the exposure and comparison groups based on information collected at baseline only. Whilst this approach was likely chosen due to inherent limitations of data sources, it did not consider the nature of depressive symptoms and clinical depression. Depressive symptom levels are known to fluctuate over time and whilst some individuals are diagnosed with clinical depression once, others are diagnosed with recurrent depressive episodes (Colman &

Ataullahjan, 2010). It is likely that the risk of cardiovascular events differs between individuals with different depressive symptom trajectories over time and between individuals with one or multiple episodes of clinical depression. It should be noted that studies using repeat assessments of the exposure differed in their methodological approach, each of which had different strengths and limitations. For example, studies defining their exposure based on the proportion of observations with depression or depressive symptoms considered that participants who were depressed once might not be depressed at other points in time. However, they also assumed an accumulation of risk that was independent of the timing of depressive symptoms. For example, a participant that was depressed at the first and second wave of data collection had the same exposure level as a participant that was exposed at the fifth and sixth exposure wave. Another approach was to define the exposure based on changes in depressive symptoms at two successive waves. The authors argued that an investigation of changes of depressive symptoms is more meaningful than an assessment of depressive symptoms at one point in time since it more closely mimics the effect a potential intervention in RCTs (Glymour & Kubzansky, 2017). Furthermore, the author emphasised that time-constant confounders were controlled for due to the use of within-person change in exposure levels. However, the approach only assessed the impact of changes in depressive symptoms in two exposure waves on the risk of outcome in the next exposure wave which might have introduced bias if the latency period linking the exposure and outcome was longer.

Fourth, assessments of depression might be influenced by other constructs such as anxiety. Depression and anxiety are highly comorbid across the life course which might be unsurprising considering the overlap in diagnostic symptoms (Zbozinek et al, 2012). Although studies using measures of general psychological distress were excluded, due to overlap in symptoms it is very likely that the exposure of included studies captured symptoms of depression as well as anxiety. The effect estimates might therefore not solely reflect the association between depression and hazards of stroke or MI but partly reflect increased hazards of stroke and MI among individuals with anxiety. It remains to be established to what extent the increased hazard among

individuals with depression or depressive symptoms relative to non-exposed individuals was due to unique symptoms of depression or shared symptoms of depression and anxiety. Some included studies adjusted their effect estimates for anxiety to prevent anxiety from acting as a confounder. However, this might have had unintended effects if the adjustment for anxiety simultaneously adjusted for the effect of shared symptoms of depression and anxiety.

Fifth, individuals with subclinical CVD might have been misclassified as depressed due to the overlap of somatic symptoms in both disorders such as chest tightness, fatigue, shortness of breath, and sleep disturbance (Kapfhammer, 2006). Barefoot & Schroll (1996) conducted a sensitivity analysis in which they excluded somatic symptoms from the depressive symptom rating scale. Whilst effect estimates were slightly lower after exclusion of somatic symptoms, CIs largely overlapped and Barefoot & Schroll (1996) concluded that results were unchanged. Other studies used the Hospital Anxiety and Depression Scale (HADS) (Gustad et al, 2014; Langvik & Hjemdal, 2015) and the Geriatric Depression Scale (GDS) (Köhler et al, 2013), two depressive symptom rating scales that do not contain questions on somatic symptoms. Gustad et al (2014) and Langvik & Hjemdal (2015) observed increased risks of MI among individuals with depression relative to non-exposed individuals. Köhler et al (2013) observed an increased risk of stroke in unadjusted but not fully adjusted model. However, effect estimates were imprecise. If individuals with subclinical CVD were indeed more likely to be misclassified as being depressed than people without subclinical CVD, this would have biased the estimates away from the null value.

3.5.2.2.2 Outcome assessment

Inaccurate measurement of the outcome might also have introduced bias. Whilst some studies solely relied on hospital and cause of death records to identify events during follow-up, others additionally used self-reported information to identify non-fatal events. Whilst the inclusion of self-reported events might have led to the inclusion of more false-positive events, the former studies will have missed non-fatal

MI and stroke events that were not treated in the hospital setting. In addition, recall and response bias might have influenced studies that relied on self-reported events. In a systematic review assessing the accuracy of electronic health records for identifying stroke events, Woodfield et al (2015) concluded that positive predictive values of more than 90% can be achieved when stroke-specific codes were chosen. This was true for stroke and each pathological subtype. McCormick et al (2014) performed a systematic review to assess the validity of electronic health records for identifying individuals with MI. Whilst equally high positive predictive values were seen for MI using hospital records (92% in more than half of all included studies), the highest positive predictive value for MI in cause of death records was 59%. Therefore, McCormick et al (2014) concluded that hospital records appear to provide valid data on MI whilst the accuracy of cause of death records was lower.

3.5.2.3 Confounding and mediation

It is important to consider the timing of covariates when differentiating between confounders and mediators. Confounders are common sources of both the exposure and outcome that preceded the exposure and outcome. Mediators are covariates that lie on the causal pathway and occur after the exposure but precede the outcome. Whilst some studies might have failed to condition on common sources of depression and cardiovascular events because of which estimates are influenced by confounding (Figure 3.14a), other studies might have conditioned on factors that lie on the causal pathway, thereby blocking causal effects and introducing bias.

Different methods of conditioning on covariates were used in included studies. Most studies used adjustment in statistical models, restriction and/ or stratification to condition on covariates. Adjustment, restriction, and stratification are different methods to condition on common sources of both the exposure and outcome, all of which are aimed at closing an open backdoor path and thereby preventing confounding (Figure 3.14b). Whilst these methods prevent confounding in situations with common sources of both the exposure and outcome, it introduces bias if covariates are on the causal pathway. Two studies applied stabilised inverse

probability weights. The aim of inverse probability weighting is to create a pseudo population in which exposed and unexposed individuals are unconditionally exchangeable (i.e. no confounding) (Hernán et al, 2000). If inverse probability weights are estimated correctly, the exposure is no longer related to measured confounders (missing arrow in Figure 3.14c). As a result it is assumed that the risk of CVD among the unexposed group would have been the same as the risk of CVD among the exposed group had individuals in the unexposed group been exposed. In contrast to adjustment, restriction, and stratification, inverse probability weighting does not block the effects of variables that lie on the causal pathway (missing arrow from covariate to exposure instead of square border in Figure 3.14c). It is important to note that covariates might act as confounding and mediating factor at the same time when time-varying exposure and covariate assessments are present (see section 7.3.2 The need for an accurate definition of depression for a detailed discussion).

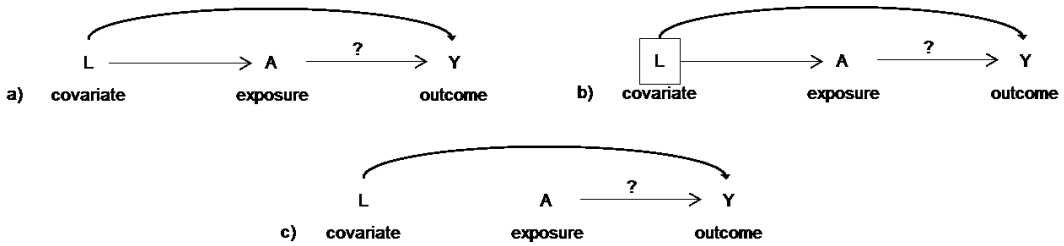


Figure 3.14: Directed acyclic graphs of confounding (a) and methods of preventing confounding: conditioning on common sources (b) and inverse probability weighting (c)
NB: Presence of arrow; causal effect is assumed or unwilling to assume that causal effect does not exist; absence of arrow: willing to assume that causal effect does not exist; direction of arrow: direction of causal effect; arrow with question mark: causal path of interest; square border: conditioning on variable through adjustment, stratification or restriction

Residual confounding has likely influenced the results of some existing studies. Residual confounding occurred when studies failed to condition on common sources of both depression and stroke or MI. Whilst most studies conditioned on sociodemographic factors, only 11 studies conditioned on smoking, physical activity, alcohol intake, and a measure of weight (e.g. BMI). Assuming that all of these risk factors preceded the onset of depression and stroke or MI, these factors would be potentially important confounding factors. Furthermore, comorbidities at baseline might have acted as confounding factors if the onset of the comorbidity was before

the onset of the exposure and outcome. Another potential reason for residual confounding was measurement error of covariates. This has likely influenced the estimates of some included studies due to recall and response bias.

Effect estimates were likely to be biased due to conditioning on mediating factors. Some authors argued that some covariates were on the causal pathway but adjusted for these covariates in the analysis (Majed et al, 2012; Moise et al, 2016; O'Brien et al, 2015; Péquignot et al, 2016). This introduced bias since a causal pathway between depression and stroke/ MI was blocked. Furthermore, many studies mentioned mediating factors in the background and discussion section but did not consider the implications of this assumption in their analysis. The exception to that are six studies that either reported that they did not condition on specific covariates since they were assumed to be on the causal pathway (Glymour et al, 2010; Glymour et al, 2012; Gustad et al, 2014; Wassertheil-Smoller et al, 2004) or that applied inverse probability weights to control for time-updated confounding without conditioning on potential mediating pathways. (Gilsanz et al, 2017; Gilsanz et al, 2015).

Overall, studies did not sufficiently justify their chosen analytical approach. For example, most studies did not discuss the importance of the timing of covariates in order to differentiate between confounders and mediators. Furthermore, studies discussed mechanisms in their discussion section but did not take these mechanisms into account in their own analysis. Also, the implications of using adjustments instead of inverse probability weighting in the presence of time-updated exposures and covariates were neither discussed nor taken into account in the majority of studies.

3.5.3 Strengths and limitations of this systematic review

This systematic review has major strengths. After screening 10,476 primary studies that have been published since January 2010 and 49 articles that were identified through screening of reference lists of existing reviews and published primary studies, this systematic review identified 51 studies that were eligible for inclusion. The large number of screened studies reflects the comprehensive search of multiple electronic databases and the development of a sensitive search strategy which

lowered the risk of missing potentially relevant studies. Furthermore, this systematic review provided an update of existing reviews on the risk of stroke or MI among individuals with depression or depressive symptoms, relative to unexposed individuals. Twelve additional primary studies were included in this meta-analysis that have not been identified in existing meta-analyses on the risk of total stroke, ischaemic stroke or haemorrhagic stroke (Cummings et al, 2016; Daskalopoulou et al, 2016; Everson-Rose et al, 2014; Gafarov et al, 2017; Gilsanz et al, 2017; Mathur et al, 2016; Mejía-Lancheros et al, 2014; Moise et al, 2016; O'Brien et al, 2015; Péquignot et al, 2016; Stewart et al, 2016; Sun et al, 2016). Moreover, 11 studies were included in this systematic review and meta-analysis that have not been included in either of the two existing meta-analyses on the risk of MI (Chi et al, 2014; Cummings et al, 2016; Daskalopoulou et al, 2016; Gafarov et al, 2017; Joyce, 2015; Mathur et al, 2016; Mejía-Lancheros et al, 2014; Péquignot et al, 2013; Stewart et al, 2016; Stürmer et al, 2006; Wassertheil-Smoller et al, 2004). Also, this is the first systematic review that considered studies which assessed depressive symptoms on continuous scales. The identification of eligible studies was thorough and transparent. A quality assessment of included studies was conducted to ensure a detailed assessment of the strengths and weaknesses of studies and identify potential sources of errors. The identification of a large number of eligible studies allowed a detailed analysis of the current evidence base. Furthermore, through pooling results of multiple primary studies in a meta-analysis, more weight was given to more robust studies with bigger sample sizes.

Some limitations of this systematic review should be noted. First, whilst the decision to look at risk of stroke and MI increased comparability between methods applied in different parts of this project, it has narrowed the picture of the literature and studies of potential interest to the overall aim of this project might have been missed. Second, the systematic review and meta-analysis was conducted by only one person. Ideally, there would have been an independent review of all search results by a second reviewer. However, due to the large number of search hits this was not deemed feasible at the time of the conduct of the review. As a result, reviewer bias might have

influenced findings since decisions were not challenged by a second reviewer. Third, although there was a protocol for the systematic review, the protocol was not published on a database of systematic reviews such as PROSPERO (Page et al, 2018). Ideally, the protocol would have been registered to ensure that researchers can compare planned and reported methods and to avoid unwanted duplication of work (Stewart et al, 2012). Fourth, the findings of this systematic review and meta-analysis were directly influenced by strengths and weaknesses of included studies. Clinical and/ or methodological differences between studies led to moderate to substantial between-study heterogeneity in the analysis on the hazard of MI among individuals with depression or depressive symptoms relative to non-exposed individuals. Subgroup analyses were performed, and sex was identified as a potential source of heterogeneity. However, as described above the lower between-study heterogeneity in the subgroup analysis of studies that provided sex-specific results might also be explained by the exclusion of studies that did not provide results separately for men and women. Between-study heterogeneity was low to moderate in the meta-analysis on the hazard of stroke. In addition, publication bias might have influenced findings. Publication bias is a source of error that results from the failure to publish studies with small effect sizes or negative findings (Dickersin & Min, 1993). If studies with negative findings or small effects have not been published, this would have biased the estimates of the meta-analysis away from null. A formal investigation of publication bias, for example using funnel plots and Egger test, was not performed as part of this systematic review since the aim of this project was to identify gaps in the research that potentially could be overcome as part of this project. There was evidence for publication bias in existing systematic reviews and meta-analyses on the risk of stroke or MI (Barlinn et al, 2014; Dong et al, 2012; Gan et al, 2014; Pan et al, 2011b; Wu & Kling, 2016). Correction for publication bias using trim-and-fill methods slightly attenuated results of existing reviews but risks of stroke and MI among individuals with depression or depressive symptoms remained higher than among unexposed individuals. Moreover, it should be noted that it is uncommon to publish protocols of analyses using observational studies (Dal-Ré et al, 2014). As a result,

findings of included studies might be the result of fishing for discoveries in datasets, and null findings might not get published. If fishing was present, this would have biased the estimates away from the null value.

3.5.4 Comparison of findings with previous systematic reviews

3.5.4.1 Stroke

The results of four existing systematic reviews and meta-analyses on the association between depression or depressive symptoms and risk of different types of stroke were consistent with the results of this systematic review and meta-analysis (Barlinn et al, 2014; Dong et al, 2012; Li et al, 2015; Pan et al, 2011b). All existing meta-analyses reported pooled effect estimates on the risk of total stroke among individuals with depression or depressive symptoms relative to non-exposed individuals (see section 2.3.2.2.2 Depression among individuals with no history of cardiovascular disease for pooled estimates). The reported associations were stronger than the associations in this meta-analysis but CIs overlapped. None of the existing meta-analyses reported the risk of total stroke per unit increase in depressive symptom scales. Studies included in existing reviews but excluded from this review were screened but used outcome definitions that were not in keeping with inclusion criteria of this review or included participants with a history of stroke. Two existing meta-analyses reported the risk of ischaemic and haemorrhagic stroke among individuals with depression or depressive symptoms relative to non-exposed individuals (Li et al, 2015; Pan et al, 2011b). In contrast to this meta-analysis, both of the existing reviews included studies in which participants with a history of stroke were not excluded. The strength of associations was very similar to the one reported in this meta-analysis (see section 2.3.2.2.2 Depression among individuals with no history of cardiovascular disease for pooled estimates).

3.5.4.2 Myocardial infarction

The results of the two existing systematic reviews and meta-analyses on the association between depression and risk of MI are consistent with the results of this systematic review and meta-analysis (Gan et al, 2014; Wu & Kling, 2016). The reported associations were slightly stronger than the association in this meta-analysis (see

section 2.3.2.2.2 Depression among individuals with no history of cardiovascular disease for pooled estimates). However, differences might be due to chance alone. Neither of the existing meta-analyses reported the risk of MI per unit increase on depressive symptom scales. Some of the studies included in existing reviews were excluded in this review since they included participants with CVD at baseline (Gump et al, 2005), reported non-comparable effect estimates (Pratt et al, 1996), were based on a selected population (Cohen et al, 2015), used a measure of psychological distress (Whang et al, 2009), or used an outcome definition that was not consistent with the inclusion criteria of this review (Capistrant et al, 2013; Hawkins et al, 2014; Huang et al, 2013; Surtees et al, 2008).

3.5.5 Implications for research

This systematic review and meta-analysis suggests an association between depression or depressive symptoms and hazards of stroke or MI. However, it also highlights shortcomings of the current evidence base which suggests that more research is needed. First, more studies with repeat assessments of the exposure are needed. It is of interest to investigate to what extent the cardiovascular risk differs between individuals with different depressive symptom trajectories and between individuals with single or recurrent episodes of clinical depression. Second, although a large number of primary studies was identified, it is not clear whether associations were similar in population subgroups. In this review, sex differences and differences between studies using different assessments of the exposure were investigated. However, there might be differences between individuals from different ethnic groups, SES, and with and without comorbidities. Only a limited number of studies performed subgroup and interaction analyses, possibly due to small sample sizes and/or number of events. Such analyses might help in further defining the population at particularly high risk of stroke and MI. Third, the role of covariates needs to be investigated further. Specifically, the timing of exposure, covariates and outcomes needs to be assessed and considered in more detail. Future studies should justify the methodological approach taken in the analysis. For example, were covariates considered to be common sources of both the exposure and outcome or were they

considered to be on the causal pathway? If covariates were assumed to be on the causal pathway, mediation analyses might be able to further our understanding of the pathways between depression and risk of stroke and MI. However, benefits and limitations of conducting mediation analyses in the context of depression should be carefully considered (see section 7.3.3 The challenge of establishing causality between depression and CVD).

3.6 Conclusion

This review aimed to identify and critically appraise existing research and identify shortcomings of existing studies that could potentially be addressed as part of this project. The meta-analysis of 51 studies suggests that the hazard of stroke and MI is higher among individuals with depression or depressive symptoms relative to non-exposed individuals. However, findings were influenced by shortcomings of existing studies. The next chapters of this thesis present data analyses of the UK Biobank (Chapter 4), the ALSWH (Chapter 5), and the Whitehall II study (Chapter 6). The analyses benefitted from strengths of each of the datasets and addressed different limitations of the prior evidence base.

Chapter 4: The association between different measures of depression and subsequent major cardiovascular events

4.1 Background

The observed association between depression or depressive symptoms and subsequent CVD might be partly explained by residual confounding. The residual confounding hypothesis originated from findings that depressed patients are markedly different from non-depressed patients in terms of their sociodemographic characteristics as well as lifestyle and biological factors (Penninx, 2016a). Consequently, these factors should be taken into account in epidemiological studies of the association between depression and CVD. Due to the strength and consistency of the association some researchers argue that the observed associations are unlikely explained by residual confounding alone (Penninx, 2016a), others emphasise the lack of adjustments in primary research (Nicholson et al, 2006) and argue that variation of estimates might be due to variation in statistical adjustment (Stapelberg et al, 2013).

Due to the large amount of data collected the UK Biobank offers a unique opportunity to control for a wide range of potentially important confounding factors. The UK Biobank is a prospective study of 500,000 participants aged between 40 and 69 years, recruited in 22 assessment centres in England, Scotland, and Wales from 2006 to 2010 (Sudlow et al, 2015). During the baseline assessment, participants completed a touchscreen questionnaire, which included questions on sociodemographic details (such as age, sex, education, income, and area-based deprivation), medical history (such as self-reported doctor diagnoses), lifestyle factors (such as smoking habit and physical activity), and medication use. In addition, participants were asked to participate in a computer-assisted verbal interview with a nurse and to undergo physical measurements such as body weight, height, and blood pressure. Health outcomes of UK Biobank participants are captured through data linkage to routine health records including primary and secondary care health records, cancer registries and death records.

A further unique advantage of the UK Biobank is its large sample size. This offers an opportunity to identify potential effect modifying factors of the risk of CVD among individuals with depression or depressive symptoms. Identifying characteristics of individuals with depression or depressive symptoms that put them at particularly high or low risk of adverse health events is important since it might allow tailoring of resources and preventive measures. Furthermore, the identification of effect-modifying factors might help inform our understanding of the association between depression and CVD from a mechanism perspective. Since depression is highly comorbid with physical disorders (Air et al, 2016), it was of particular interest to assess to what extent the association between depression and CVD differs between individuals with and without comorbidities (see section 2.3.1 The relationship between major mental disorders and physical disorders for more detailed discussion of mental-physical comorbidity). Depression has been shown to be associated with worse self-management strategies and treatment adherence among individuals with physical comorbidities (DiMatteo et al, 2000; Gonzalez et al, 2008; Sjösten et al, 2013). Therefore, individuals with depression who additionally have a diagnosis of physical health conditions might be at particularly high risk of adverse health outcomes, such as MCVE. In addition, mental-physical comorbidity was shown to be more prevalent among individuals with lower socioeconomic background (Barnett et al, 2012; McLean et al, 2014). It is of interest to assess to what extent individuals with lower socioeconomic background not only have higher absolute risks of mental-physical comorbidity but also have a more pronounced relative risk of cardiovascular events associated with depression or depressive symptoms than individuals from a higher socioeconomic background.

This analysis aimed to explore the association between different measures of depression and MCVE in the UK Biobank whilst taking account of a wide range of potential confounding factors. Furthermore, the role of comorbidities (hypertension, diabetes, cholesterol levels) and socioeconomic factors (education, area-based deprivation) as potential effect-modifying factors was investigated.

4.2 Objectives

1. To explore the association between different measures of depression and MCVE, defined as first-ever fatal or non-fatal stroke and first-ever fatal or non-fatal MI, as well as stroke and MI separately
2. To investigate the role of comorbidities (hypertension, diabetes, cholesterol levels) and socioeconomic factors (education, area-based deprivation) as potential effect-modifying factors

4.3 Methods

4.3.1 Sample

The study population included UK Biobank participants without a history of CVD and major mental disorders other than depression. History of CVD and major mental disorders other than depression were defined as present if a participant reported a history of any of the disorders in the touchscreen questionnaire or nurse interview, or if an event was identified through linkage to hospital inpatient records in England, Scotland, and Wales (Table 4.1). History of CVD was defined as history of stroke, MI, angina or TIA before baseline. In accordance with UK Biobank recommendations, stroke was defined as subarachnoid haemorrhage, intracerebral haemorrhage, cerebral infarction, or unspecified stroke (ICD-10: I60.X, I62.X, I64.X, I64.X) and MI was defined as acute or subsequent MI, Dressler syndrome, and old MI (ICD-10: I21.X, I22.X, I24.1, I25.2) (Schnier et al, 2017a; Schnier et al, 2017b). TIA was defined using ICD-10 codes G45.0, G45.1, G45.8 and G45.9 and angina using I20.0, I20.1, I20.8, and I20.9. History of major mental disorders other than depression was defined as history of bipolar disorder or schizophrenia (ICD-10: F20.X, F21, F22.X, F23.X, F24, F25.X, F28, F29, F30.X, F31.X). In contrast to UK Biobank recommendations, prevalent events were not identified using ICD-9 codes in order to improve the consistency of look-back periods for events in Scotland, Wales and England. Participants whose records could not be linked with hospital inpatient or cause of death data, with missing or conflicting information in the cause of death dataset, with unreasonable death dates (date prior to attending the assessment centre), and individuals who have since withdrawn from the UK Biobank were excluded from the analysis.

Table 4.1: Overview of available information on hospital inpatient data (UK Biobank, 2018)

Data source	Data provider	ICD-9	ICD-10	Period of data currently available	Censoring date
Hospital Episode Statistics for England	NHS Digital, England	---	1996 onwards	1996 onwards	31 March 2015
Scottish Morbidity Record	ISD, Scotland	1981 to 1996	1996 onwards	1997 onwards	31 October 2015
Patient Episode Database for Wales	SAIL, Wales	---	1999 onwards	1998 onwards	29 February 2016

ISD: Information and Statistics Division, NHS: National Health Services, SAIL: Secure Anonymised Information Linkage

The analyses of UK Biobank data were conducted under generic approval from the NHS National Research Ethics Service (Ref 11/NW/0382, approval letter dated 17 June 2011). Full written informed consent was gained from participants at the point of data collection. This research has been conducted using the UK Biobank Resource under application number 13797.

4.3.2 Definition of variables

4.3.2.1 Exposure

The exposures of interest were different measures of depression (Table 4.2). Antidepressant use and self-reported doctor diagnosis of depression were identified in a nurse interview at baseline. Antidepressant use was defined as self-reported use of at least one selective serotonin reuptake inhibitor (SSRI) or 'other' antidepressant in accordance with the list described by Martin et al (2016). Hospital diagnosis with depression was defined as a hospital inpatient record with depressive episode or recurrent depressive episode (ICD-10: F32.X, F33.X) before baseline. Depression was defined as at least one of antidepressant use, self-reported depression, or hospital diagnosis with depression at baseline. In the following, the term depression is used when referring to the composite definition of depression.

Table 4.2: Overview of the definition of different measures of depression

Exposure	Definition
Depression	At least one of antidepressant use, hospital diagnosis with depression or self-reported doctor diagnosis of depression
Antidepressant use	Self-reported use of at least one of the following antidepressants: Prozac 20mg capsule, sertraline, lustral 50mg tablet, paroxetine, fluoxetine, fluvoxamine, seroxat 20mg tablet, citalopram, cipramil 10mg tablet, escitalopram, cipralex 5mg tablet, Limbitrol 10 capsule, prothiaden 25mg capsule, elavil 10mg tablet, lentizol 25mg m/r capsule, tryptizol 10mg tablet, anafranil 10mg capsule, tofranil 10mg tablet, lofepramine, trimipramine, surmontil 10mg tablet, bolvidon 10mg tablet, norval 10mg tablet, phenelzine, nardil 15mg tablet, moclobemide, manerix 150mg tablet, limbitrol-5 capsule, triptafen tablet, amitriptyline+chlordiazepoxide 12.5mg/5mg capsule, amitriptyline, hydrochloride+perphenazine 10mg/2mg tablet, mianserin, amitriptyline, clomipramine, dothiepin, imipramine, gamanil 70mg tablet, dosulepin, maoi- phenelzine, venlafaxine, efexor 37.5mg tablet, lomont 70mg/5ml s/f suspension, mirtazapine, zispin 30mg tablet, thaden 25mg capsule, duloxetine, yentreve 20mg gastro-resistant capsule, cymbalta 30mg gastro-resistant capsule
Self-reported depression	Self-reported doctor diagnosis of depression in nurse interview at baseline
Hospital diagnosis with depression	Records on depressive episode or recurrent depressive episode (ICD-10: F32.X, F33.X) in the hospital episode statistics before baseline

4.3.2.2 Outcome

MCVEs were defined as first-ever fatal or non-fatal stroke or MI, whichever occurred first. Non-fatal stroke and MI during follow-up were identified through linkage to hospital inpatient data (Table 4.1). The ICD codes used to identify events during follow-up were consistent with the definition used to identify prevalent events (see section 4.3.1 Sample). In order to increase comparability between participants from England, Scotland, and Wales, the same censoring date was used for all participants (31 March 2015). Fatal stroke and MI were identified through linkage to cause of death registers (Table 4.3). Cause and date of death were provided by NHS Digital for participants from England and Wales, and by Information Services Division for participants from Scotland. Again, the same censoring date was chosen for all participants (30 November 2015). Survival times were calculated from the date of attending the baseline assessment centre to the date of first-ever MCVEs, date of death, or end of follow-up.

Table 4.3: Overview of available information on date and cause of death (UK Biobank, 2018)

Data source	Data provider	ICD-9	ICD-10	Period of data currently available	
				Start	End
Death register England & Wales	NHS Digital	---	2006 onwards	April 2006 onwards	31 January 2016
Death register Scotland	ISD	---	2006 onwards	April 2006 onwards	30 November 2015

ISD: Information and Statistics Division, NHS: National Health Services

4.3.2.3 Covariates

The following covariates were ascertained through self-report on the touchscreen questionnaire at baseline. Income was defined as average household income before tax and categorised into five groups (less than £18,000, £18,000 to 30,999, £31,000 to 51,999, £52,000 to 100,000, greater than £100,000). Education was defined as highest educational attainment and categorised into three groups (College or university degree; A levels, O levels, Certificate of Secondary Education (CSE) or National Vocational Qualification (NVQ), or equivalent; and none of the above). Area-based deprivation was assessed through the Townsend Deprivation Index (Townsend, 2009), and divided into fifths within the study population. The Townsend index was recalculated for all UK Biobank participants using the 2001 UK Census (UK Biobank Data Management Team, 2017). Due to small numbers of participants with a non-white ethnicity, ethnicity was categorised as white or other ethnicity. Smoking was categorised as never, previous, or current smoker. Alcohol was defined as safe drinking if men or women had ≤ 14 units of alcohol per week and as risky drinking if alcohol intake exceeded 14 units per week which is in accordance with the weekly intake guideline of the UK Department of Health (UK Chief Medical Officer, 2016). Physical activity was categorised as low, moderate or high level of physical activity in accordance with the International Physical Activity Questionnaire (IPAQ Group, 2010). BMI was calculated based on height and weight measurements ascertained at the assessment centre at baseline. Due to small numbers of participants with underweight BMI was categorised as under- or normal weight (< 18.5 – 24.9 kg/m^2), overweight (25 – 29.9 kg/m^2), obese (30 – 34.9 kg/m^2), severely obese (35 – 39.9 kg/m^2), morbidly obese (> 40 kg/m^2). Oily fish intake was categorised into three categories (at

least once a week, less than once per week, never). Following a recommended intake approach, total fruit and vegetable intake was treated as binary variable assessing whether or not participants consumed at least five fruits or vegetables a day (World Health Organisation and Food and Agriculture Organisation of the United Nations, 2004). Family history of stroke, heart disease, high blood pressure, and severe depression, was defined as self-reported illness of mother or father. Diabetes, hypertension, and high cholesterol were defined as diagnosis and/ or treatment, ascertained through self-report in the touchscreen questionnaire or nurse interview. Hypertension was additionally ascertained through blood pressure measurements of ≥140/90mm Hg at the assessment centre at baseline.

4.3.3 Statistical analyses

4.3.3.1 Association between depression and subsequent MCVE, stroke, and MI

Baseline characteristics were assessed separately for participants with and without the exposures and outcomes of interest. Due to the large sample size no formal tests for differences between groups were performed. Cox proportional hazards models were performed to estimate HRs and 95% CI of the hazard of MCVE among participants with depression, antidepressant use, hospital admission with depression, and self-reported depression relative to participants without the exposure of interest. In accordance with the residual confounding hypothesis, measured covariates were treated as potential confounding factors of both depression and MCVE. Implicitly the assumption was made that all covariates measured at baseline preceded depression and CVD onset. The first Cox proportional hazards model represented the unadjusted association between different measures of depression and MCVE. The second model was adjusted for age, sex, ethnicity, and all socioeconomic factors (education, income, and area-based deprivation). The third model additionally controlled for BMI, alcohol intake, physical activity, smoking, fruit and vegetable consumption, oily fish intake, family history of CVD, and family history of depression. Since age did not fulfil the assumption of a linear contribution to the Cox proportional hazards model, it was split into four equally sized groups and added as categorical variable to the model. The proportional hazard assumption

was checked visually using log-minus-log survival plots and by testing the proportional hazard assumption (cox.zph function). There was no evidence of a violation of the proportional hazard assumption in any of the fully adjusted models. The results of the global tests of the proportional hazard assumption of each fully adjusted model are presented in Appendix Table A 9. Furthermore, the results of the checks for proportionality for all covariates of one of the models are presented as an example (Appendix Table A 10). In secondary analysis, the association between depression and first-ever fatal or non-fatal stroke and first-ever fatal or non-fatal MI were analysed separately. Since participants were censored if they died from causes other than stroke or MI, it was of interest to investigate to what extent deaths from causes other than stroke and MI might affect associations between depression and MCVE by running a competing risk analysis. Therefore, cumulative incidence functions of fatal or non-fatal stroke or MI and death from causes other than stroke or MI among participants with and without any indication of depression at baseline were described. In addition, cause-specific HRs (95% CI) of the association between depression and deaths from causes other than stroke or MI were calculated.

4.3.3.2 Multiplicative and/ or additive interaction

Evidence of multiplicative and/ or additive interaction between depression and socioeconomic factors or comorbidities was investigated in order to determine to what extent these factors modify the effect of depression on MCVE. The presence of multiplicative interaction was tested by adding a product term of the two variables of interest to the fully adjusted Cox proportional hazards model. The null hypothesis of no multiplicative interaction between depression and the second variable of interest was rejected at the 5% significance level ($p < 0.05$). The presence of additive interaction was investigated by calculating the relative excess risk due to interaction (RERI), the attributable proportion due to interaction (AP), and the synergy index. All measures of additive interaction and their corresponding 95% CIs were derived in keeping with Andersson et al (2005). A RERI estimate close to zero indicates that there is little excess risk due to interaction. Similarly, an AP close to zero indicates that the proportion of events due to interaction is small. The synergy index shows the excess

risk from the exposure to both factors when interaction is present relative to the risk from being exposed to both factors when interaction is absent. Therefore, a synergy index of one indicates that interaction is absent. Educational attainment and area-based deprivation were treated as binary variables in the interaction analyses between depression and socioeconomic factors. Educational attainment was defined as high for participants with a university or college degree and as low for participants with no university or college degree. Area-based deprivation was defined as high if the Townsend score had a positive value, whereas it was defined as low if the Townsend score had a negative value. A Townsend score of 0 indicated an area in the UK with mean values. A score of 0 did not occur in the sample.

4.3.3.3 Multiple imputation

The patterns of missing data as well as the plausibility of different mechanisms of missingness were investigated. The pattern of missing data was investigated by describing the number and proportion of missing data for each variable and participant. Baseline characteristics of participants with and without any missing data were compared in order to explore to what extent missing data are missing completely at random (MCAR). Differences between participants with and without complete data were indicative of a violation of the MCAR assumption. Since a MCAR mechanism was deemed unlikely, univariate and multivariate logistic regressions were performed in order to determine predictors of missingness. A missing data indicator variable was used as dependent variable in these models. The multivariate logistic regressions were adjusted for age, sex, ethnicity, education, income, and area-based deprivation. Since a missing at random (MAR) mechanism was deemed likely, multiple imputation of missing data was performed using the MICE package in R (van Buuren & Groothuis-Oudshoorn, 2011). One imputation for each percent of incomplete cases in the dataset is recommended (van Buuren, 2012). In keeping with that, 40 imputations with ten iterations were performed. Since it is essential that the imputation model is in keeping with the substantive model of the analysis, imputations were performed separately for each exposure-outcome combination. In addition, interactions of interest had to be added to the imputation model. Since all

interaction terms were interactions between two binary variables, one of which was fully observed (depression), the imputations were run separately for participants with and without depression. Thus, the association between the second variable of the interaction term and cardiovascular events could differ by depression status (van Buuren, 2012). Since a complete case analysis is biased when missing data are not MCAR, the analyses presented in this chapter are based on imputed data. The latter were compared to results obtained through complete-case analyses.

4.4 Results

4.4.1 Study population

Out of 502,655 UK Biobank participants, 465,215 participants were included in the analysis (Figure 4.1).

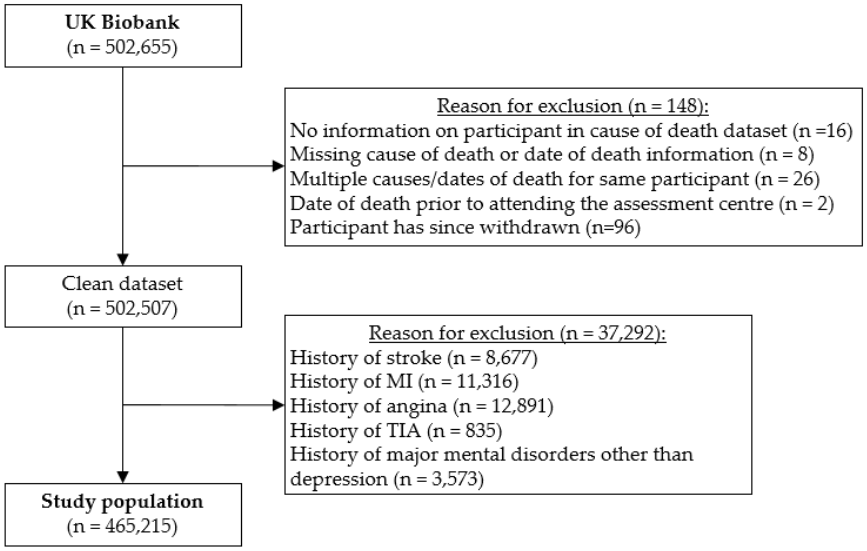


Figure 4.1: Flow chart of UK Biobank participants included in analyses

4.4.2 Investigation of missing data mechanism

4.4.2.1 Missingness pattern

The average number of variables with missing information was low (0.6 variables per participant) but overall missingness was high due to the large number of variables of interest (Table 4.4). Out of all participants 39.3% had at least one missing value in any of the variables. Among participants with at least one variable with missing information, the largest proportion had missing information on one variable (25.9%). Missing values were present in 12 out of 27 variables of interest (Table 4.5). The

frequency of missingness was below 5.0% for all variables except income (15.1%), alcohol intake (14.8%), family history of depression (13.0%), and family history of CVD (7.6%).

Table 4.4: Number (%) of missing variables per participant

Number of missing variables	Number of participants (%)
0	282,190 (60.7)
1	120,659 (25.9)
2	43,410 (9.3)
3	13,757 (2.0)
4	3,324 (0.7)
5	725 (0.2)
6	158 (< 0.1)
7	49 (< 0.1)
8	37 (< 0.1)
9	31 (< 0.1)
10	538 (< 0.2)
11	335 (< 0.1)
12	2 (< 0.01)

Table 4.5: Number (%) of missing values in each variable of interest

Variable	Number missing (%)
Income	70,079 (15.1)
Alcohol intake	68,817 (14.8)
Family history of depression	60,595 (13.0)
Family history of cardiovascular disease	35,497 (7.6)
Physical activity	18,961 (4.1)
Highest educational attainment	9,144 (2.0)
Oily fish intake	3,641 (0.8)
Body mass index	2,661 (0.6)
Smoking	2,599 (0.6)
Ethnicity	2,523 (0.5)
Fruit and vegetable intake	1,534 (0.3)
Area-based deprivation	575 (0.1)
Depression, age, sex, hypertension, diabetes, cholesterol, major cardiovascular events, stroke, myocardial infarction	0 (0)

4.4.2.2 Plausibility of different missingness mechanisms

The plausibility of a MCAR mechanism was investigated by comparing baseline characteristics of participants with and without complete data available (Table 4.6). Differences in baseline characteristics were observed for all variables. However, differences in self-reported depression, smoking, fruit and vegetable intake, oily fish intake, and family history of depression were less pronounced than differences among the other variables of interest. The proportion of participants with any measure of depression, obesity, low physical activity levels, hypertension, diabetes, or high cholesterol levels was lower among participants with complete data available than among participants with at least one missing variable. In contrast, the proportion of male participants and participants with a college or university degree and high income was higher among complete cases.

*Table 4.6: Baseline characteristics separately for participants with and without complete data available (page 1 of 2)**

	Incomplete case (n = 183,025, 39.3%)	**Complete case** (n = 282,190, 60.7%)
Depression (%)	17,969 (9.8)	22,384 (7.9)
Antidepressant use (%)	13,997 (7.6)	17,133 (6.1)
Self-reported depression (%)	10,785 (5.9)	14,354 (5.1)
Hospital diagnosis with depression (%)	1,574 (0.9)	1,657 (0.6)
Male (%)	72,453 (39.6)	132,925 (47.1)
Age (median [IQR])	58.0 [50.0, 64.0]	57.0 [49.0, 62.0]
Ethnicity = Other ethnic groups (%)	13,565 (7.5)	11,222 (4.0)
Income (%)		
Greater than 100,000	3,142 (2.8)	19,115 (6.8)
52,000 to 100,000	16,728 (14.8)	66,534 (23.6)
31,000 to 51,999	27,501 (24.3)	77,839 (27.6)
18,000 to 30,999	31,339 (27.7)	68,623 (24.3)
Less than 18,000	34,236 (30.3)	50,079 (17.7)
Highest educational attainment (%)		
College or university degree	42,290 (24.3)	110,852 (39.3)
A levels, O levels, CSE, NVQ, or equivalent	89,565 (51.5)	139,488 (49.4)
None of the above	42,026 (24.2)	31,850 (11.3)
Area-based deprivation (%)		
1 (Least deprived)	33,826 (18.5)	61,125 (21.7)
2	34,394 (18.9)	59,334 (21.0)
3	35,670 (19.6)	57,943 (20.5)
4	36,086 (19.8)	56,752 (20.1)
5 (Most deprived)	42,474 (23.3)	47,036 (16.7)
Body mass index (%)		
Under- or normal weight	56,840 (31.5)	100,932 (35.8)
Overweight	74,823 (41.5)	121,549 (43.1)
Obese	33,706 (18.7)	44,443 (15.7)
Severely obese	10,512 (5.8)	11,354 (4.0)
Morbidly obese	4,483 (2.5)	3,912 (1.4)
Physical activity (%)		
High	13,675 (8.3)	28,652 (10.2)
Moderate	109,883 (67.0)	194,877 (69.1)
Low	40,506 (24.7)	58,661 (20.8)
Alcohol intake = Risky drinking (%)	48,255 (42.3)	139,654 (49.5)
Smoking status (%)		
Never	101,991 (56.5)	156,822 (55.6)
Previous	57,396 (31.8)	98,755 (35.0)
Current	21,039 (11.7)	26,613 (9.4)

*Table 4.6 continued: Baseline characteristics separately for participants with and without complete data available (page 2 of 2)**

	Incomplete case (n = 183,025, 39.3%)	**Complete case** (n = 282,190, 60.7%)
Fruit and vegetable intake per day = Less than five a day (%)	124,832 (68.8)	198,211 (70.2)
Oily fish intake (%)		
At least once a week	96,833 (54.0)	159,652 (56.6)
Less than once per week	59,279 (33.0)	95,226 (33.7)
Never	23,272 (13.0)	27,312 (9.7)
Hypertension (%)	104,016 (56.8)	148,647 (52.7)
Diabetes (%)	9,936 (5.4)	10,158 (3.6)
High cholesterol levels (%)	29,967 (16.4)	38,823 (13.8)
Family history of cardiovascular disease (%)	113,503 (76.9)	206,631 (73.2)
Family history of depression (%)	12,706 (10.4)	28,398 (10.1)

* Data are given as n (%) unless specified

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

In addition, differences in the proportion of individuals with a MCVE, stroke, and MI and the time of follow-up were investigated among complete and incomplete cases (Table 4.7). The proportion of individuals with a MCVE, stroke and MI was lower among cases with complete data than among cases with incomplete data. Also, the median follow-up time was 0.3 years shorter among cases with complete data than among individuals with at least one missing variable for all outcomes of interest.

*Table 4.7: Differences in events and follow-up times by complete case status**

	Incomplete case (n = 183,025, 39.3%)	**Complete case** (n = 282,190, 60.7%)
Major cardiovascular events (%)	3,648 (2.0)	4,201 (1.5)
Years of follow-up (median [IQR])	7.0 [6.3, 7.6]	6.7 [6.0, 7.4]
Stroke (%)	1,452 (0.8)	1,673 (0.6)
Years of follow-up (median [IQR])	7.0 [6.3, 7.6]	6.7 [6.0, 7.4]
Myocardial infarction (%)	2,269 (1.2)	2,616 (0.9)
Years of follow-up (median [IQR])	7.0 [6.3, 7.6]	6.7 [6.0, 7.4]

* Data are given as n (%) unless specified

Since the observed differences between cases with complete and incomplete data indicate a violation of the MCAR assumption, the plausibility of a MAR mechanism was investigated by performing univariate and multivariate logistic regression

analyses with a missing data indicator as dependent variable (Table 4.8). All variables of interest were predictors of missingness in univariate and multivariate logistic regression models. The odds of having at least one variable with missing data was higher among participants with depression, females, an ethnic background other than white, lower SES, higher BMI, lower physical activity level, less oily fish intake, hypertension, diabetes, high cholesterol, and a family history of CVD or depression. The odds of having at least one variable with missing values was lower among individuals in the second, third and fourth quartiles relative to the youngest quartile, and among participants with risky drinking behaviour.

*Table 4.8: Odds ratios (95% CI) of univariate and multivariate logistic regression analyses with a missing data indicator as dependent variable (page 1 of 2)**

Variable	Unadjusted OR (95% CI)	Adjusted OR* (95% CI)
Depression (Any vs. none)	1.26 (1.24 – 1.29)	1.15 (1.12 – 1.18)
Antidepressant use (Yes vs. no)	1.28 (1.25 – 1.31)	1.14 (1.10 – 1.17)
Self–reported depression (Yes vs. no)	1.17 (1.14 – 1.20)	1.12 (1.09 – 1.16)
Hospital diagnosis with depression (Yes vs. no)	1.47 (1.37 – 1.57)	1.13 (1.04 – 1.23)
Sex (Female vs. Male)	1.36 (1.34 – 1.38)	1.12 (1.11 – 1.14)
Ethnicity (Other vs. white ethnic group)	1.96 (1.91 – 2.01)	1.51 (1.46 – 1.56)
Age, quartiles		
1 (youngest)	ref.	ref.
2	0.97 (0.96 – 0.99)	0.86 (0.84 – 0.88)
3	1.15 (1.14 – 1.17)	0.79 (0.78 – 0.81)
4 (oldest)	1.51 (1.48 – 1.53)	0.83 (0.81 – 0.85)
Income		
Greater than 100,000	ref.	ref.
52,000 to 100,000	1.53 (1.47 – 1.59)	1.48 (1.42 – 1.54)
31,000 to 51,999	2.15 (2.07 – 2.24)	1.96 (1.88 – 2.04)
18,000 to 30,999	2.78 (2.67 – 2.89)	2.38 (2.28 – 2.48)
Less than 18,000	4.16 (4.00 – 4.33)	3.07 (2.94 – 3.21)
Highest educational attainment		
College or university degree	ref.	ref.
A levels, O levels, CSE, NVQ, or equivalent	1.68 (1.66 – 1.71)	1.32 (1.30 – 1.34)
None of the above	3.46 (3.40 – 3.52)	1.95 (1.90 – 2.00)
Area–based deprivation, fifths		
1 (Least deprived)	ref.	ref.
2	1.05 (1.03 – 1.07)	1.00 (0.98 – 1.02)
3	1.11 (1.09 – 1.13)	1.04 (1.01 – 1.06)
4	1.15 (1.13 – 1.17)	1.05 (1.02 – 1.07)
5 (Most deprived)	1.63 (1.60 – 1.66)	1.29 (1.26 – 1.33)

*Table 4.8 continued: Odds ratios (95% CI) of univariate and multivariate logistic regression analyses with a missing data indicator as dependent variable (page 2 of 2)**

Variable	Unadjusted OR (95% CI)	Adjusted OR* (95% CI)
Body mass index		
Under– or normal weight	ref.	ref.
Overweight	1.09 (1.08 – 1.11)	1.11 (1.10 – 1.13)
Obese	1.35 (1.32 – 1.37)	1.31 (1.29 – 1.34)
Severely obese	1.64 (1.60 – 1.69)	1.56 (1.51 – 1.62)
Morbidly obese	2.03 (1.95 – 2.13)	1.91 (1.81 – 2.01)
Physical activity		
High	ref.	ref.
Moderate	1.18 (1.16 – 1.21)	1.03 (1.01 – 1.06)
Low	1.45 (1.41 – 1.48)	1.23 (1.19 – 1.26)
Alcohol intake (Risky vs. safe drinking)	0.75 (0.74 – 0.76)	0.95 (0.93 – 0.97)
Smoking		
Never	ref.	ref.
Previous	0.89 (0.88 – 0.91)	0.87 (0.86 – 0.88)
Current	1.22 (1.19 – 1.24)	1.08 (1.05 – 1.10)
Fruit and vegetable intake per day (Less than five a day vs. at least five a day)	0.93 (0.92 – 0.95)	0.98 (0.96 – 0.99)
Oily fish intake		
At least once a week	ref.	ref.
Less than once per week	1.03 (1.01 – 1.04)	1.13 (1.12 – 1.15)
Never	1.40 (1.38 – 1.43)	1.33 (1.30 – 1.36)
Hypertension (Yes vs. no)	1.18 (1.17 – 1.20)	1.04 (1.03 – 1.06)
Diabetes (Yes vs. no)	1.54 (1.49 – 1.58)	1.35 (1.30 – 1.39)
High cholesterol levels (Yes vs. no)	1.23 (1.21 – 1.25)	1.08 (1.05 – 1.10)
Family history of cardiovascular disease (Yes vs. no)	1.22 (1.20 – 1.24)	1.33 (1.31 – 1.36)
Family history of depression (Yes vs. no)	1.03 (1.01 – 1.06)	1.12 (1.09 – 1.15)

* adjusted for age, sex, ethnicity, education, income, and area–based deprivation

ref.: reference, CSE: Certificate of secondary education, NVQ: National vocational qualification

All unadjusted and adjusted Cox proportional hazard models indicated increased hazards of MCVE, stroke, and MI among individuals with at least one variable with missing values relative to individuals with complete data (Table 4.9).

*Table 4.9: Hazard ratios (95% CI) of MCVE, stroke, and MI among individuals with at least one variable with missing information relative to complete cases**

	Unadjusted hazard ratios (95% CI)	Adjusted hazard ratios* (95% CI)
Major cardiovascular events	1.30 (1.25 – 1.36)	1.20 (1.14 – 1.26)
Stroke	1.30 (1.22 – 1.40)	1.13 (1.04 – 1.23)
Myocardial infarction	1.30 (1.23 – 1.37)	1.23 (1.15 – 1.32)

* adjusted for age, sex, ethnicity, education, income, and area–based deprivation

One cannot distinguish between MAR and missing not at random (MNAR) mechanisms since MNAR mechanisms depend on data that were not collected. However, a MAR mechanism was deemed likely given the large number of variables in the dataset and the numerous significant associations between observed variables and missingness. In the following, it is therefore assumed that data are independent of unobserved variables, given all observed variables. Multiple imputation was performed since a complete case analysis is likely biased when the MCAR assumption is violated. Therefore, the following results are based on imputed data. The results of a complete case analysis are presented in Appendix Table A 27 to Table A 34 and differences between the complete cases analysis and the analysis based on imputed data are discussed in section 4.4.5 Comparison of results based on complete cases and imputed data. The convergence of all multiple imputations was checked and deemed sufficient. The convergence plots of one exposure-outcome set are presented in Appendix Table A 11 as an example.

4.4.3 To what extent are different measures of depression associated with subsequent major cardiovascular events, stroke, and myocardial infarction?

4.4.3.1 Descriptive statistics

Among 465,215 UK Biobank participants included in the analysis, 205,378 (44.1%) were male, and the median age was 57 (IQR: 50 – 63) years. The vast majority of participants had a white ethnic background (94.1%). A total of 40,353 (8.7%) participants were defined as depressed at baseline, of whom 25,139 (62.3%) had a self-reported depression diagnosis, 3,231 (8.0%) had a hospital admission with depression, and 31,130 (77.1%) reported using antidepressants (Figure 4.2). 17,766 (44.0%) of all participants with depression were identified through more than one data source, whereas 13,692 (33.9%) were identified through antidepressant use alone, 7,765 (19.2%) through self-report alone, and 1,130 (2.8%) through hospital diagnosis with depression alone.

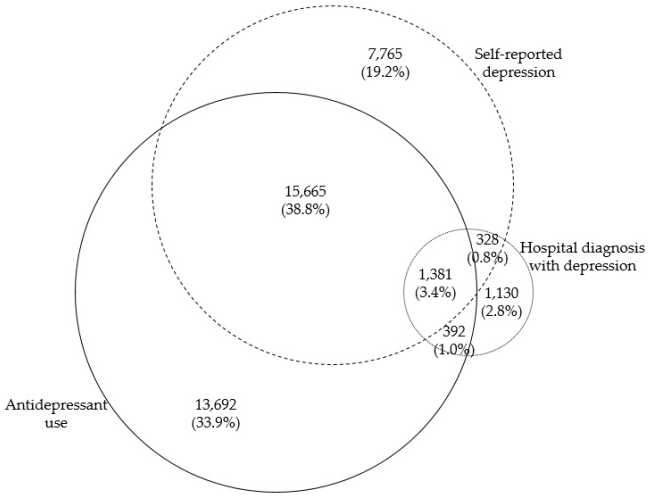


Figure 4.2: Venn diagram representing the relationship between each measure of depression

There were differences between participants with and without depression at baseline (Table 4.10). The proportion of male participants was lower among participants with versus without depression (30.4% and 45.5%, respectively). Furthermore, the proportion of participants with an average household income of more than £31,000, with a college or university degree, and in the three least deprived fifths was lower among participants with versus without depression. The proportion of individuals with obesity, who were previous or current smokers, physically inactive, or reported having diabetes or high cholesterol levels was higher whereas risky drinking behaviour was lower among participants with depression compared to those without depression. The proportion of participants with hypertension was almost identical among participants with and without depression at baseline (54.4% and 54.2%, respectively). Baseline characteristics separately for participants with and without antidepressant use, hospital diagnosis with depression, self-reported depression, and participants with and without MCVE, MI, and stroke during-follow-up are presented in Appendix Table A 12 to Table A 17.

*Table 4.10: Baseline characteristics separately for participants with and without one or more of self-reported depression, antidepressant use, or hospital admission with depression (page 1 of 2)**

	No depression (n = 424,862, 91.3%)	**Depression** (n = 40,353, 8.7%)
Male (%)	193,121 (45.5)	12,257 (30.4)
Age (median [IQR])	57.0 [50.0, 63.0]	57.0 [50.0, 62.0]
Ethnicity (%)		
White	399,138 (93.9)	38,767 (96.1)
Other ethnic groups	23,398 (5.5)	1,389 (3.4)
Income (%)		
Greater than 100,000	21,308 (5.0)	949 (2.4)
52,000 to 100,000	78,691 (18.5)	4,571 (11.3)
31,000 to 51,999	97,593 (23.0)	7,747 (19.2)
18,000 to 30,999	90,983 (21.4)	8,979 (22.3)
Less than 18,000	72,522 (17.1)	11,793 (29.2)
Highest educational attainment (%)		
College or university degree	141,857 (33.4)	11,285 (28.0)
A levels, O levels, CSE, NVQ, or equivalent	208,854 (49.2)	20,199 (50.1)
None of the above	65,732 (15.5)	8,144 (20.2)
Area-based deprivation (%)		
1 (Least deprived)	87,999 (20.7)	6,952 (17.2)
2	86,466 (20.4)	7,262 (18.0)
3	85,983 (20.2)	7,630 (18.9)
4	84,549 (19.9)	8,289 (20.5)
5 (Most deprived)	79,362 (18.7)	10,148 (25.1)
Body mass index (%)		
Under- or normal weight	146,201 (34.4)	11,571 (28.7)
Overweight	180,520 (42.5)	15,852 (39.3)
Obese	69,968 (16.5)	8,181 (20.3)
Severely obese	18,882 (4.4)	2,984 (7.4)
Morbidly obese	6,908 (1.6)	1,487 (3.7)
Physical activity (%)		
High	39,596 (9.3)	2,731 (6.8)
Moderate	280,569 (66.0)	24,191 (59.9)
Low	87,724 (20.6)	11,443 (28.4)
Alcohol intake = Risky drinking (%)	174,947 (41.2)	12,962 (32.1)
Smoking status (%)		
Never	238,884 (56.2)	19,929 (49.4)
Previous	142,258 (33.5)	13,893 (34.4)
Current	41,337 (9.7)	6,315 (15.6)
Fruit and vegetable intake per day = Less than five a day (%)	295,180 (69.5)	27,863 (69.0)
Oily fish intake (%)		
At least once a week	235,774 (55.5)	20,711 (51.3)
Less than once per week	140,871 (33.2)	13,634 (33.8)
Never	44,927 (10.6)	5,657 (14.0)

*Table 4.10 continued: Baseline characteristics separately for participants with and without one or more of self-reported depression, antidepressant use, or hospital admission for depression (page 2 of 2)**

	No depression (n = 424,862, 91.3%)	**Depression** (n = 40,353, 8.7%)
Hypertension (%)	230,315 (54.2)	22,348 (55.4)
Diabetes (%)	17,609 (4.1)	2,485 (6.2)
High cholesterol levels (%)	61,055 (14.4)	7,735 (19.2)
Family history of cardiovascular disease (%)	291,388 (68.6)	28,746 (71.2)
Family history of depression (%)	33,709 (7.9)	7,395 (18.3)

* This table does not show the number and percentage of missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

4.4.3.2 The association between different measure of depression and major cardiovascular events, stroke, and myocardial infarction

During a median of 6.8 (IQR: 6.1 – 7.5) years of follow-up, 7,849 MCVE occurred. There were increased hazards of MCVE among individuals with versus without depression in all models, irrespective of the different measures of depression (Figure 4.3, Appendix Table A 18). The strength of the association between depression, antidepressant use, and self-reported depression and MCVE increased after adjustment for age, sex, ethnicity, and socioeconomic factors and attenuated after adjustment for comorbidities, lifestyle factors, and family history of CVD and depression. The strength of the association between hospital diagnosis with depression and MCVE attenuated with each adjustment. There was a 21% increased hazard of MCVE among individuals with depression relative to participants without depression in the fully adjusted model (HR: 1.21, 95% CI: 1.12 – 1.30). The association was strongest among individuals with a hospital admission with depression, followed by participants with antidepressant use, and individuals with self-reported depression but CIs overlapped (Figure 4.3, Appendix Table A 18).

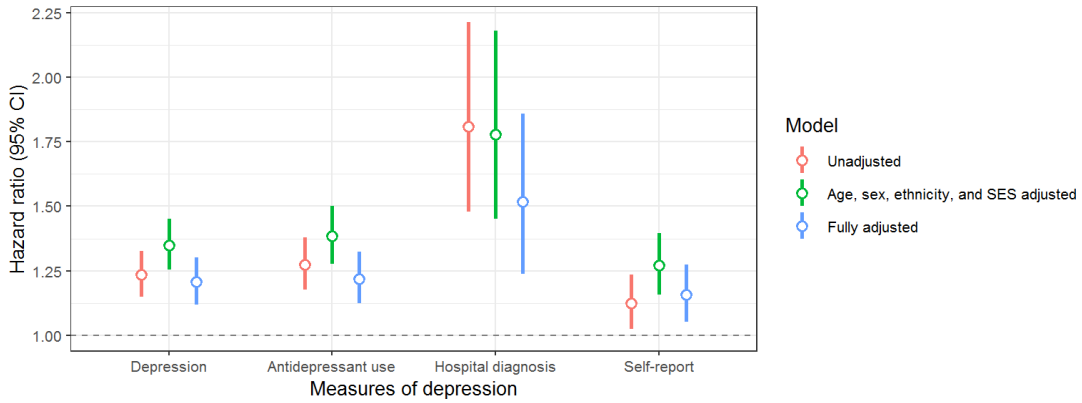


Figure 4.3: Hazard ratios (95% CI) of the association between different measures of depression and MCVE in unadjusted, partially adjusted and fully adjusted models (imputed data)

Similar patterns were observed when stroke and MI were used separately as outcomes. In total, there were 4,885 MI and 3,125 strokes during follow-up. There was an 18% increased hazard of stroke and 24% increased hazard of MI among individuals with depression, relative to unexposed individuals, in fully adjusted models (HR, 95% CI: 1.18, 1.05 – 1.33, and 1.24, 1.12 – 1.36, respectively). Whilst point estimates of the association between depression and MI were slightly higher than the point estimate of the association between depression and stroke, the pattern of the associations was similar for both outcomes and CIs overlapped (Appendix Table A 19).

Depression was associated with increased hazards of MCVE among both men and women (fully adjusted HR, 95% CI: 1.12, 1.00 – 1.24 and 1.31, 1.17 – 1.45, respectively), but the association was significantly stronger among women (p for multiplicative interaction: 0.01). The observed sex difference was largely driven by a statistically significant interaction between depression and sex in the analyses on MI but not stroke risk (p for multiplicative interaction: <0.01 and 0.55, respectively). On an additive scale, the hazards of MCVE were higher among women with depression, men without depression, and men with depression relative to women without depression (fully adjusted HR, 95% CI: 1.33, 1.20 – 1.47; 2.35, 2.23 – 2.49, and 2.60, 2.33 – 2.90, respectively) (Figure 4.4). The hazard of MCVE among men with depression

was approximately equivalent to the sum of the hazards due to male gender and depression alone.

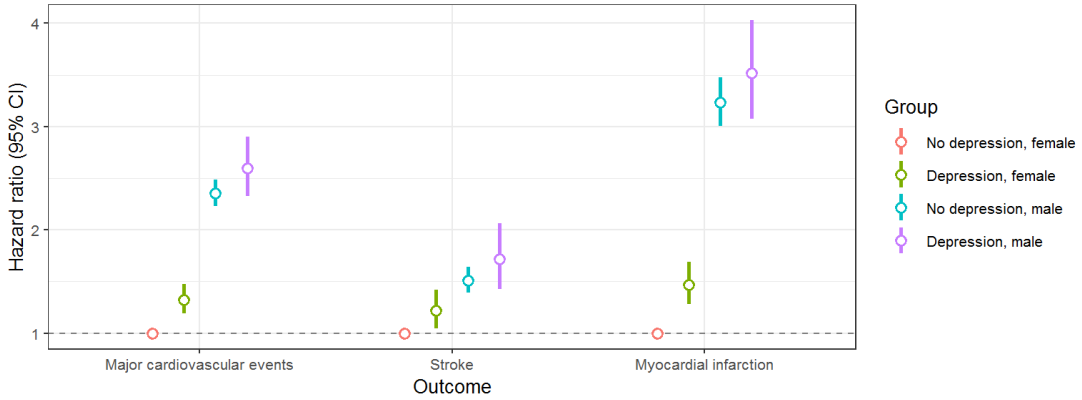


Figure 4.4: Hazard ratio (95% CI) of major cardiovascular events, stroke, and myocardial infarction among women with depression, men without depression and men with depression, relative to women without depression (imputed data)

4.4.3.3 Competing risk analysis

The probability of dying from causes other than stroke or MI was higher among participants with depression than among participants without depression throughout follow-up (Figure 4.5). Furthermore, the cause-specific Cox proportional hazards model indicated increased hazards of deaths from causes other than MVCE among participants with depression, relative to participants without depression in unadjusted, partially adjusted and fully adjusted models (HR, 95% CI: 1.40, 1.32 – 1.49; 1.40, 1.32 – 1.49, and 1.30, 1.22 – 1.39, respectively).

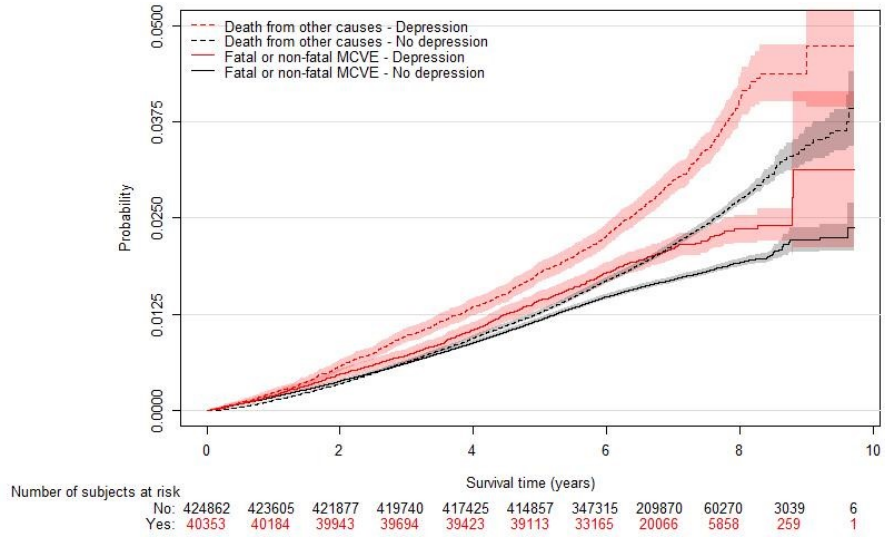


Figure 4.5: Cumulative incidence function of fatal or non-fatal stroke or MI and death from causes other than stroke or MI among participants with versus without any indication of depression at baseline

4.4.4 Do comorbidities or socioeconomic factors modify the association between depression and major cardiovascular events?

4.4.4.1 Comorbidities

Information on depression, hypertension, diabetes, and cholesterol levels was used to create three four-level predictor variables indicating whether a participant had neither depression nor the comorbidity of interest, depression alone, the comorbidity of interest alone, or both depression and the comorbidity of interest at baseline (Table 4.11). In the analyses on depression and/ or diabetes, and depression and/ or high cholesterol levels, the largest proportion of participants was allocated to the group with neither depression nor the comorbidity of interest at baseline (87.5% and 78.2%, respectively). In contrast, in the depression and/ or hypertension analysis, the hypertension alone category was most common (49.5%), followed by the group with neither depression nor hypertension at baseline (41.8%). Similarly, whilst the neither depression nor the comorbidity of interest category was least common in the analyses on depression and/ or diabetes and depression and/ or high cholesterol analyses (0.5% and 1.7%, respectively), the depression alone category was least common in the analyses on depression and/ or hypertension (3.8%).

*Table 4.11: Distribution of the four level predictor variables depression and/ or hypertension, depression and/ or diabetes, and depression and/ or high cholesterol levels**

	No hypertension	**Hypertension**
No depression	194,547 (41.8%)	230,315 (49.5%)
Depression	18,005 (3.9%)	22,348 (4.8%)
	No diabetes	**Diabetes**
No depression	407,253 (87.5%)	17,609 (3.9%)
Depression	37,868 (8.1%)	2,485 (0.5%)
	Low cholesterol levels	**High cholesterol levels**
No depression	363,807 (78.2%)	61,055 (13.1%)
Depression	32,618 (7.0%)	7,735 (1.7%)

* Data are given as n(%)

There were differences between the groups with depression and/ or hypertension, depression and/ or diabetes, and depression and/ or high cholesterol levels at baseline (Appendix Table A 20 to Table A 22). The proportion of male participants was highest

among the groups with hypertension alone, high cholesterol levels alone, and diabetes alone and lowest among the groups with depression alone. The median age was higher among the groups with hypertension, diabetes, or high cholesterol levels than among the groups without the comorbidity of interest, irrespective of whether or not depression was present. The groups with comorbid depression and hypertension, diabetes, or high cholesterol levels were most likely to live in the most deprived area, to report less than £18,000 of income, and were least likely to have a college or university degree. The groups with both depression and at least one of the comorbidities had worse cardiovascular risk profiles at baseline in terms of the proportion of individuals with severe or morbid obesity, low physical activity, current smoking, oily fish intake, and were more likely to report a family history of CVD and depression. The groups with depression alone showed similar characteristics with regard to current smoking and oily fish intake. No clear pattern was observed for risky drinking behaviour.

Individuals with depression and at least one of hypertension, diabetes, and high cholesterol levels were all at particularly high hazard of MCVE. However, the pattern of the association between the four level predictor variables and MCVE was different for each of the comorbidities of interest (Figure 4.6, Appendix Table A 23). Hypertension alone and comorbid depression and hypertension were associated with greater hazards of MCVE relative to participants with neither depression nor hypertension (1.78, 95% CI: 1.68 – 1.89, and 2.23, 95% CI: 2.03 – 2.45, respectively). In contrast, there was no increased hazard of MCVE among the group with depression alone relative to the group with neither depression nor hypertension (1.06, 95% CI: 0.91 – 1.24). In the analysis on depression and/ or diabetes, there was a gradual increase in hazard of MCVE from individuals with depression alone (1.22, 95% CI: 1.13 – 1.32) to participants with diabetes alone (1.47, 95% CI: 1.35 – 1.61), to participants with depression and diabetes (1.63, 95% CI: 1.33 – 2.00), relative to participants with neither depression nor diabetes. In the analysis on depression and/ or high cholesterol levels, depression alone and comorbid depression and high cholesterol levels were associated with greater hazards of MCVE relative to

participants with neither depression nor high cholesterol levels in all models (1.17, 95% CI: 1.07 – 1.28, and 1.24, 95% CI: 1.09 – 1.41, respectively). High cholesterol levels alone were associated with an increased hazard of MCVE in the unadjusted and partially adjusted models (1.90, 95% CI: 1.80 – 2.01, and 1.22, 95% CI: 1.15 – 1.29, respectively) but not in the fully adjusted model (0.97, 95% CI: 0.91 – 1.03).

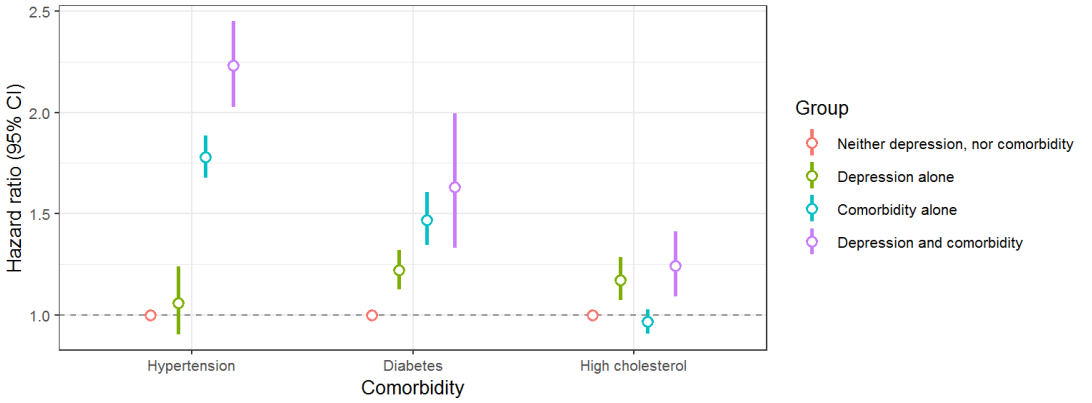


Figure 4.6: Hazard ratios (95% CI) of the association between depression and/ or hypertension, depression and/ or diabetes, and depression and/ or high cholesterol levels and MCVE in fully adjusted models (imputed data)

Whilst there was formal evidence for additive interaction between depression and hypertension, evidence for interaction neither was found between depression and diabetes nor between depression and high cholesterol levels (Table 4.12). There was an excess risk due to interaction between depression and hypertension on the additive scale (RERI: 0.39, 95% CI: 0.16 – 0.63), with 17.5% (95% CI: 7.5 – 27.5%) of all MCVE events estimated to be due to the interaction between depression and hypertension (attributable proportion). In keeping with that, the excess risk from being exposed to depression and hypertension when interaction is present, relative to the risk when interaction is absent was 1.47 (95% CI: 1.14 – 1.89). There was no evidence for additive interaction between depression and diabetes and depression and high cholesterol levels. However, the CIs of all of the measures were wide indicating imprecise estimation. None of the product terms between depression and hypertension, depression and diabetes, and depression and high cholesterol levels reached the conventional threshold for statistical significance indicating that one

should not reject the null hypothesis of no multiplicative interaction between depression and each of the comorbidities.

*Table 4.12: Results of formal tests for additive and multiplicative interaction between depression and hypertension, depression and diabetes, and depression and high cholesterol levels**

	RERI	Attributable proportion	Synergy index	Multiplicative interaction
Depression and hypertension	0.39 (0.15 – 0.63)	17.5 (7.5 – 27.5)	1.47 (1.14 – 1.89)	p = 0.06
Depression and diabetes	-0.06 (-0.41 – 0.29)	3.6 (-25.8 – 18.5)	0.91 (0.53 – 1.58)	p = 0.40
Depression and high cholesterol levels	0.10 (-0.09 – 0.29)	8.1 (-6.2 – 22.3)	1.71 (0.62 – 4.76)	p = 0.26

* Data are presented as estimate (95% confidence interval) unless specified

RERI: Relative excess risk for interaction

4.4.4.2 Socioeconomic factors

Two four level predictor variables were created indicating the presence of depression and/ or low educational attainment and depression and/ or high area-based deprivation (Table 4.13). Out of the four potential depression and educational attainment combinations, low educational attainment alone was most common (60.2%) and depression alone was least common (2.5%). Out of the four potential depression and area-based deprivation combinations, neither depression nor high area-based deprivation was most common (66.2%) and depression and high area-based deprivation (3.0%) was least common.

*Table 4.13: Distribution of the four level predictor variables depression and/ or low educational attainment and depression and/ or high area-based deprivation**

	High educational attainment	Low educational attainment
No depression	141,857 (31.1%)	274,586 (60.2%)
Depression	11,285 (2.5%)	28,343 (6.2%)
	Low area-based deprivation	**High area-based deprivation**
No depression	307,806 (66.2%)	116,553 (25.1%)
Depression	26,318 (5.7%)	13,963 (3.0%)

* Data are presented as n (%)

There were differences between the groups with depression and/ or low education, and depression and/ or high area-based deprivation at baseline (Appendix Table A

24 and Table A 25). The proportion of male participants was higher among the groups without depression than among the groups with depression. The proportion of individuals with an income less than £18,000, severe and morbid obesity, diabetes, high cholesterol levels, and who were current smokers was highest among the groups with depression and low educational attainment and depression and high area-based deprivation. The proportion of individuals with low physical activity levels was highest among the groups with depression, irrespective of the socioeconomic factor. Risky drinking behaviour was most prevalent among individuals without depression and high educational attainment or low area-based deprivation. Whilst there was little difference in the proportion of individuals with hypertension between the groups with depression and/ or high area-based deprivation, the proportion of individuals with hypertension was about 10% higher among the groups with low educational attainment than among the groups with high educational attainment.

Individuals with depression and low educational attainment and individuals with depression and high area-based deprivation showed the highest hazards of MCVE relative to the group with no depression and high educational attainment/ low area-based deprivation (Figure 4.7, Appendix Table A 26). Depression alone, high area-based deprivation alone, and combined depression and high area-based deprivation were associated with a higher hazard of MCVE relative to participants with no depression and low area-based deprivation (1.10, 95% CI: 1.00 – 1.22, 1.07, 95% CI: 1.02 – 1.13, and 1.46, 95% CI: 1.31 – 1.63). Similarly, depression alone, low educational attainment alone, and combined depression and low educational attainment were associated with a higher hazard of MCVE relative to participants with no depression and high educational attainment (1.16, 95% CI: 0.99 – 1.37, 1.11, 95% CI: 1.05 – 1.18, and 1.35, 95% CI: 1.23 – 1.49).

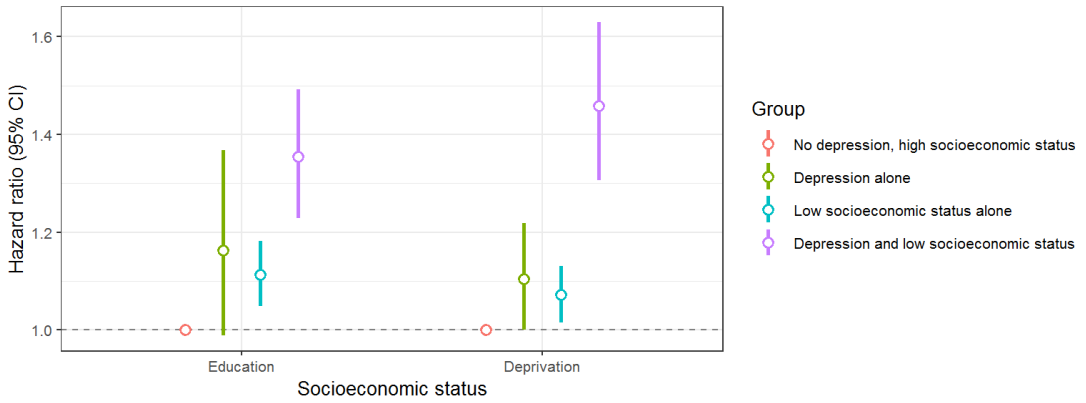


Figure 4.7: Hazard ratios (95% CI) of the association between depression and/ or low educational attainment and depression and/or high area-based deprivation in fully adjusted models (imputed data)

Although the pattern of the associations with MCVE looked similar for both four level predictor variables, evidence for additive and multiplicative interaction was found between depression and area-based deprivation but not between depression and educational attainment (Table 4.14). The product term between depression and high area-based deprivation reached the conventional threshold for statistical significance whereas the product term between depression and low educational attainment did not ($p < 0.01$, and $p = 0.63$, respectively). The RERI between depression and high area-based deprivation was estimated at 0.28 (95% CI: 0.10 – 0.47), with an estimated 19.4% (95% CI: 8.0 – 30.8%) of MCVE that was due to interaction between depression and area-based deprivation. The excess risk from being exposed to depression and area-based deprivation when interaction is present, relative to the risk when interaction is absent was 2.61 (95% CI: 1.22 – 5.60). The RERI between depression and low educational attainment was estimated at 0.08 (95% CI: -0.14 to 0.29). The proportion of MCVE that was due to interaction between depression and low educational attainment was estimated at 5.8% (95% CI: -10.0 to 21.5%). The excess risk from being exposed to depression and low educational attainment when interaction is present, relative to the risk when interaction is absent was 1.28 (95% CI: 0.60 – 2.72). The CIs of some of the estimates of additive interaction were wide indicating imprecise estimation.

*Table 4.14: Results of formal tests for additive and multiplicative interaction between depression and educational attainment and depression and area-based deprivation**

	RERI	Attributable proportion	Synergy index	Multiplicative interaction
Depression and educational attainment	0.08 (-0.14 – 0.29)	5.8 (-10.0 – 21.5)	1.28 (0.60 – 2.72)	p = 0.63
Depression and area-based deprivation	0.28 (0.10 – 0.47)	19.4 (8.0 – 30.8)	2.61 (1.22 –5.60)	p < 0.01

* Data are presented as estimate (95% confidence interval) unless specified

RERI: Relative excess risk for interaction

4.4.5 Comparison of results based on complete cases and imputed data

The results of the complete cases analysis provided similar results as the analysis based on imputed data (Appendix Table A 27 to Table A 34). As expected, the CIs of all estimates based on imputed data were narrower than the CIs based on complete cases due to the larger sample size. Whilst the point estimates of the association between different measures of depression and MCVE were slightly higher in the analyses based on imputed data, the CIs overlapped with the estimates of the complete case analysis indicating that the observed differences might be due to chance alone. Whilst the association between depression and stroke was stronger than the association between depression and MI in all models based on complete cases, the association between depression and MI was stronger than the association between depression and stroke in all models based in imputed data. However, in both analyses CIs overlapped. The estimates of the competing risk analysis and the analyses on potential effect-modifying factors were similar in analyses based on complete cases and imputed data.

4.5 Discussion

4.5.1 Summary of key findings

In a large prospective study, depression, antidepressant use, hospital diagnosis with depression, and self-reported depression were associated with increased hazards of MCVE, even after adjustment for a wide range of potential confounding factors. In the fully adjusted model, having at least one of antidepressant use, hospital diagnosis with depression, or self-reported depression was associated with a 21% increased

hazard of MCVE relative to the hazard among participants who did not have any of these indications of depression at baseline. There was a potential dose-response relationship since the association was more pronounced between a hospital diagnosis with depression and MCVE and slightly weaker between self-reported depression and MCVE. Depression was more strongly associated with hazard of MCVE among women but this was largely driven by sex differences in the association between depression and MI but not stroke risk. Furthermore, depression was associated with increased hazard of death from causes other than stroke or MI, indicating that the observed results were unlikely explained by competing risks.

The observed association between depression and MCVE differed between individuals with and without comorbidities at baseline as well as between individuals from different socioeconomic backgrounds. Individuals with depression and hypertension, depression and diabetes, and depression and high cholesterol levels were all at particularly high hazard of MCVE although formal evidence of additive interaction was only present between depression and hypertension. No evidence of multiplicative interaction between depression and comorbidities was found in analyses based on imputed data. Similarly, individuals with depression and low educational attainment and individuals with depression and high area-based deprivation were at particularly high hazard of MCVE. Formal evidence of multiplicative and additive interaction was found between depression and area-based deprivation only. The highest hazard of MCVE was observed among individuals with depression and hypertension at baseline, relative to individuals with no depression and no hypertension.

4.5.2 Strengths and limitations of this analysis

4.5.2.1 Chance

It is reasonable to assume that the analysis was sufficiently powered where statistically significant findings that are clinically relevant were found. However, limited statistical power might have played a role in findings that were not statistically significant or imprecise. Considering the large sample size of the study,

it is unlikely that limited statistical power has influenced the results of the primary analyses whereas analyses on potential effect-modifying factors could have been affected by limited power. The sample size of this analysis was bigger than the sample sizes of 48 out of 51 existing studies on the relationship between depression and risk of MCVE (see Table 3.1). The sample size of the study by Sun et al (2016) was similar to the sample size of this analysis. A comparison of the sizes of effects that would be detectable in datasets of these sizes is complicated due to the use of different measures of depression. However, the study by Sun et al (2016) was sufficiently powered to detect interactions between depression and age and between depression and smoking status which might indicate that studies with this sample size are sufficiently powered to detect interaction effects. Nonetheless, limited statistical power might have played a role in findings that were not statistically significant.

First, the possibility that findings might be due to chance was considered by providing 95% CIs around all estimates. Due to the large sample size of the UK Biobank, the CIs around most estimates were narrow indicating precise estimation. Furthermore, the CIs did not overlap with the null value indicating that the increased hazards among individuals with depression were unlikely to be due to chance alone. Differences in the strength of the association between the different measures of depression and major CVD were observed which might reflect a dose-response relationship. Individuals with a hospital diagnosis with depression might suffer from more severe depression than individuals who were identified through self-reported depression alone. However, CIs overlapped suggesting that the observed differences might be due to chance alone. Due to smaller group sizes of the four exposure groups in analyses on potential effect-modifying factors, CIs were slightly wider indicating less precise estimation. Furthermore, in some of these analyses the CIs of the four exposure groups overlapped suggesting that the observed differences between the exposure groups might be due to chance alone.

Second, in some analyses hypotheses tests were performed to detect statistically significant differences between groups. It was decided to qualitatively discuss

differences in baseline characteristics between groups instead of performing hypothesis tests to detect statistically significant differences. The reason for that was that hypothesis tests tend to show statistically significant differences between groups in very large studies that do not necessarily indicate meaningful differences in practice. P-values were provided in analyses assessing the presence of multiplicative interaction between depression and socioeconomic factors and depression and comorbidities. Some of the p-values did not reach the conventional threshold for statistical significance ($p < 0.05$). Therefore, the null hypothesis of no interaction between these factors was not rejected. A p-value greater than 0.05 could be an indication that there truly was no interaction between these factors, but it might also be that the analyses were not sufficiently powered to detect a true difference between the groups. Other p-values reached the conventional threshold for statistical significance ($p < 0.05$). Since multiple hypothesis tests have been performed, the chance of making any type I error is greater than 5%, and observed significant results could therefore also be due to chance.

4.5.2.2 Bias

4.5.2.2.1 Selection bias

Selection bias might have influenced the results of the analyses. To avoid selection bias the study population should represent the population of interest. The population of interest were adults living in the UK, without a history of stroke, MI, angina, TIA, and mental disorders other than depression at baseline. The UK Biobank had a low response rate (5.5%) which resulted in a very healthy cohort from a high socioeconomic background (Allen et al, 2012). Allen et al (2012) acknowledged that this is problematic when estimating the prevalence or incidence of disorders. Some researchers have argued that it is less likely to influence estimates of associations between diseases given that there are large numbers of participants with different levels of risk factors in the sample (Allen et al, 2012). Nonetheless, selection bias might have influenced some results of this analysis. UK Biobank participants with a history of the previously mentioned cardiovascular disorders were excluded from the analyses if they reported a history of any of the disorders or if they had a hospital

diagnosis at baseline. Whilst it is likely that most participants with a history of these CVDs were identified through this approach, individuals who did not self-report any of these disorders and who were treated in primary care might not have been identified and excluded. Since the health care system is free of charge for anyone living in the UK, it is unlikely that the identification of participants with a history of hospitalised CVD was differential between the exposure and comparison group. However, there might have been selection bias if individuals with depression do not report history of non-hospitalised CVD as accurately as people without depression.

The comparison group should be comparable to the exposure group in all factors other than the exposure of interest. Baseline characteristics of participants were presented separately for individuals with and without the exposure of interest at baseline and observed differences between groups were discussed. Differences between groups were taken into account by restricting the analyses to participants free from CVD and by adjusting the analyses for all factors with differences between the exposure and comparison group at baseline. Multiple imputation of missing data was performed in order to overcome the potential bias of performing a complete case analysis when the MCAR assumption is violated. Missing data were assumed to be MAR, implicitly assuming that missing data did not depend on any unobserved variable, given the observed variables. Given the large range of variables collected, this assumption seemed sensible. However, it could not be formally tested.

4.5.2.2.2 Information bias

4.5.2.2.2.1 Exposure assessment

It is likely that information bias has influenced findings due to inaccurate measurement of the exposure. First, a major limitation of the analyses is that depression was only assessed once at baseline. Since depression is known to fluctuate over time, participants allocated to the exposure group might be mistakenly assumed to be depressed over the whole period of follow-up, and individuals allocated to the comparison group may develop depression during follow-up. It would have been of interest to assess depression on multiple occasions over time, but the UK Biobank did

not provide sufficient information to assess depression as time-varying exposure. There was some evidence that an assessment of depression on multiple occasions over time was warranted when looking at the proportion of people without depression at baseline who had a hospital diagnosis with depression during follow-up. As expected, the proportion of individuals with a hospital diagnosis with depression during follow-up was larger among individuals with baseline depression (n = 4,587, 11.4%) than among individuals without baseline depression (n = 3,900, 0.9%). Of the 3,900 individuals without baseline depression but with a hospital diagnosis with depression 249 individuals had a MCVE during follow-up. Whilst the diagnosis with depression among the 3,651 individuals without MCVE indicated a change in their exposure status throughout follow-up, the diagnosis with depression among the 249 individuals with MCVE during follow-up might have occurred after stroke or MI (reverse causation). In addition, it is possible that some individuals in the comparison group had depression at some point in time but neither reported depression at the baseline assessment nor had a hospital diagnosis with depression before baseline. If depression is indeed associated with an increased hazard of MCVE, then this will bias the estimates towards null.

Second, an assessment of depression once at baseline might not accurately reflect the hazard of cardiovascular events among all individuals with depression but rather represent an average risk across disparate depression subgroups. The exposure group likely represented a heterogeneous group of participants, some of whom have suffered from depression for an extended time period whilst others have had a single episode of depression or depressive symptoms. It is likely that the strength of the association between depression and cardiovascular events differed between the different subgroups of individuals with depression. This further highlights the need to use repeat assessments of depression and assess depression trajectories over time.

Third, underreporting and under-ascertainment likely introduced bias. Whilst hospital diagnosis with depression was identified through linkage to hospital records, antidepressant use, and self-reported depression was based on self-report.

Assessing depression based on self-report likely introduced recall and response bias. Recall bias will have occurred when participants did not recall whether or not they have been diagnosed with depression. Response bias will have occurred when participants chose not to provide accurate information on their exposure status. This might have happened due the stigma attached to depression or because individuals might not have considered an episode of depression in the past as important, for example because it happened long before baseline. Considering the stigma attached to a depression diagnosis, response bias likely had a bigger impact on the accuracy of the depression measure than recall bias. Ideally, primary care data would have been used to triangulate the assessment of depression. However, the UK Biobank did not offer the opportunity to link to primary care data at the time this project was undertaken. Furthermore, the use of a depressive symptom rating scale might have identified individuals with high depressive symptoms but undiagnosed or subthreshold depression. Since no measure of depressive symptoms was used in this analysis, the results might be influenced by under-ascertainment. If depressive symptoms are associated with increased hazard of MCVE, this would have biased the results towards null.

A major strength of the analysis is that different measures of depression were used. The importance of using different data sources is highlighted in the fact that 56% of all participants were identified either through antidepressant use alone (33.9%), self-reported depression alone (19.2%), or hospital diagnosis with depression alone (2.8%). If any of these measures had not been used, these individuals would not have been identified as being depressed. However, the group of participants with antidepressant use alone deserves special attention. The reason for that is that antidepressants are not only used to treat depression but also to treat medical conditions such as anxiety, insomnia, and irritable bowel syndrome (Wong et al, 2017). Therefore, an unknown proportion of the group with antidepressant use might be misclassified as depressed due to the use of antidepressants to treat other medical conditions. Conducting an analysis based on electronic medical records and a prescribing system in Canada, Wong et al (2016) estimated that 29.4% of all

antidepressants were prescribed off-label. Since the indication of the use of antidepressants was not recorded in the UK Biobank, it was not possible to differentiate between individuals who took antidepressants to treat depression and individuals who took antidepressants to treat other conditions. Misclassification of those using antidepressants for other reasons would have biased a true association between depression and MCVE towards null.

4.5.2.2.2.2 Outcome assessment

Inaccurate measurement of the outcome might have introduced bias. A major strength of the study is that participants were followed prospectively and that there was very limited loss to follow-up due to the assessment of the outcome through linked health records. Furthermore, relying on information of routinely collected health records prevented response and recall bias. Woodfield et al (2015) performed a systematic review to assess the accuracy of hospital and mortality records for identifying stroke cases. They concluded that positive predictive values were greater than 70% in most studies that used stroke-specific codes and greater than 90% in some studies. This was true for stroke and each pathological subtype. This study used ICD codes to detect stroke outcomes that were recommended by Woodfield et al (2015). However, these authors only used stroke codes in either the primary or secondary diagnostic position whereas this analysis allowed the stroke code to be in any diagnostic position. Whilst the use of any diagnostic position to identify stroke cases might have increased the number of false positive events, it is unlikely that this bias is differential between individuals with and without depression. McCormick et al (2014) performed a systematic review to assess the validity of electronic health records for identifying individuals with MI. Hospital records showed a positive predictive value of more than 92% in more than half of all included studies and the highest positive predictive value in cause of death records was 59%. Therefore, McCormick et al (2014) concluded that hospital records appear to provide valid data on MI whilst the accuracy of cause of death records was lower. A disadvantage of the existing review is that only three studies reported the accuracy of ICD-10 codes to diagnose MI. Therefore, more recent studies or an updated systematic review are

needed to comment on the accuracy of the MI codes in this analysis. The number of false-negatives was lowered by using hospital and death records to identify stroke and MI during follow-up. Whilst record linkage to hospital and cause of death records offered many benefits, non-hospitalised events that were non-fatal were missed. However, it remains to be established to what extent the inclusion of primary care data will improve accuracy of outcome ascertainment (Woodfield et al, 2015).

The outcome assessment might be biased since death censoring occurred eight months after hospitalisation censoring. This bias is non-differential since non-fatal events during this time period were missed for participants with and without the exposure of interest. If depression was indeed associated with increased hazards of fatal and non-fatal MCVE, this would have biased estimates towards null.

4.5.2.3 Confounding

Residual confounding due to unmeasured confounders or inaccurate measurement of measured confounders might have introduced bias. A major advantage of this analysis is that the UK Biobank provided information on a wide range of factors that have been discussed as potential confounding factors in the literature. The analyses controlled for confounding by demographic factors, socioeconomic factors, comorbidities, behavioural factors, and family history of CVD and depression by adjusting the analyses for these factors. In addition, the analyses were restricted to participants free from CVD at baseline. With the exception of blood pressure measurements, all measured covariates were assessed based on self-report in a touchscreen questionnaire and/ or nurse interview. Self-report of health behaviours might be inaccurate due to stigma attached to behaviours such as smoking or risky alcohol intake and due to recall bias of, for example, the amount of physical activity (Connor Gorber et al, 2009; Prince et al, 2008). Furthermore, unmeasured covariates, such as inflammatory markers or genetic factors, might have influenced findings. However, given the wide range of confounding factors considered, it is unlikely that residual confounding due to unmeasured covariates alone explained the observed associations.

Since the UK Biobank did not provide sufficient information to treat depression and covariates as time-varying factors, it was difficult to establish the temporal relationship. As a result, the assumption that all measured covariates were common sources rather than mediating factors in the relationship between depression and CVD might have been violated. In this analysis, the estimates of the fully adjusted model were adjusted for demographic, socioeconomic, and lifestyle factors, as well as comorbidities and family history of depression and CVD. If the assumption was valid that all of these covariates were common sources of both depression and cardiovascular events, then an open backdoor path was closed by conditioning on all of these factors (see section 3.4.2.5 Conditioning on covariates for more detail). Whilst it can be established with certainty that genetically determined factors such as sex or ethnicity preceded the onset of both depression and cardiovascular events, lifestyle factors and comorbidities might have occurred after the onset of depression. As a result, they might act as mediating factors rather than common sources in the relationship between depression and cardiovascular events. As shown in the directed acyclic graph (Figure 4.8), conditioning on a mediator closes a causal path between depression and cardiovascular events. If the covariates indeed acted as mediating factors in the relationship between depression and cardiovascular events, then the estimates of the fully adjusted models would represent an approximation of the direct effect rather than the total effect of depression on cardiovascular events. Since the direct effect of an exposure on an outcome is a representation of the effect of the exposure on the outcome through unknown or unmeasured mediators, an estimation of the direct effect is not meaningful in practice. Instead, an estimation of the total effect is more meaningful in practice. Under the assumption that lifestyle factors and comorbidities were mediating factors in the relationship between depression and cardiovascular events, the effect estimates of the partially adjusted models might be a closer approximation of the total effect of depression on cardiovascular events. Since the role of the covariates cannot be established with certainty in analyses without time-updated information, this problem again highlights the need for an analysis of

the association between depression and cardiovascular events in a dataset with multiple assessments of depression and covariates.

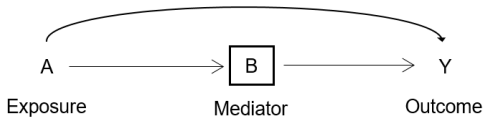


Figure 4.8: Conditioning on mediator (violation of assumption that all covariates were common sources)
NB: Presence of arrow; causal effect is assumed or unwilling to assume that causal effect does not exist; direction of arrow: direction of causal effect; square border: conditioning on variable through adjustment, stratification or restriction

4.5.3 Comparisons of findings with previous research

4.5.3.1 Existing Biobank publications

Martin et al (2016) investigated the cross-sectional association between depression and cardiometabolic diseases in UK Biobank participants with complete data of lifetime features of mood disorders and self-reported cardiometabolic diseases (n = 145,991). Depression was associated with increased risks of stroke, MI, diabetes, hypertension, and angina as well as a composite outcome of these outcomes. In a fully adjusted model adjusting for age, gender, deprivation, ethnicity, BMI, smoking status, alcohol consumption and psychotropic medication the OR was estimated at 1.26 (95% CI: 1.13 – 1.40) for stroke and at 1.18 (95% CI: 1.08 – 1.30) for MI. The strength of the association between depression and each of the outcomes was very similar in the cross-sectional and prospective analyses. Whilst the association was slightly stronger between depression and stroke than between depression and MI in the analyses by Martin et al (2016), in the current analysis the association between depression and MI was slightly stronger than the association between depression and stroke. However, the CIs of the stroke and MI estimates overlapped in both analyses. The exposure measure used by Martin et al (2016) differed from the depression measures used in this analysis. Martin et al (2016) defined depression according to a set of criteria proposed by the UK Biobank's mental health working group which was published by Smith et al (2013). A disadvantage of the criteria proposed by the UK Biobank mental health working-group is that they did not validate the criteria against established depression criteria. Furthermore, the sample size of the study by Martin

et al (2016) is much smaller than the sample size of the current analysis because additional questions on depressive and manic symptoms were necessary to define their exposure measure. These questions were added to the UK Biobank assessment at a later point in time and among only a subset of the UK Biobank sample. Another difference is that Martin et al (2016) adjusted for psychotropic medication use in the fully adjusted model. This might be problematic since depression increases the risk of taking psychotropic medication. Therefore, psychotropic medication might be seen as a mediator in the relationship between depression and CVD or as a proxy measure of the exposure itself (see section 7.3.4.1 The need to disentangle the role of depression and psychotropic medications for more detailed discussion).

The current analysis showed the same trend as an analysis investigating the prospective association between psychological distress and all-cause and cause-specific mortality (Batty et al, 2016). Psychological distress, measured by the Patient Health Questionnaire (PHQ), was associated with greater hazards of all-cause mortality and deaths due to CVD, cancer, external causes and cause-specific mortality among 308,721 UK Biobank participants. In the fully adjusted model, each one SD increase on the PHQ was associated with an increased hazard of all-cause mortality of 1.16 (95% CI: 1.12 to 1.20) and an increased hazard of CVD mortality of 1.11 (95% CI: 1.03 to 1.20). Due to the use of a continuous scale it was not possible to compare the strength of the association to the current analysis.

4.5.3.2 The association between depression and cardiovascular events

The estimates of the current analysis were very similar to the pooled estimates of the systematic review that was performed as part of this project (see Chapter 3) and slightly lower than the pooled point estimates of the most recently published meta-analyses (Li et al, 2015; Wu & Kling, 2016). However, CIs overlapped indicating that observed differences might be due to chance or variation in statistical adjustment. Only few existing studies reported results separately for men and women (see Chapter 3 for more detailed discussion). Among these studies, depression or depressive symptoms were associated with increased risks of stroke and MI among

both men and women with little difference in the strength of the association. In keeping with existing studies, depression was associated with increased risk of MCVE among both men and women. In contrast to existing studies, the association was significantly stronger among women than men, which was largely driven by a statistically significant interaction between depression and sex in the analyses on MI but not stroke risk.

4.5.3.3 The role of socioeconomic factors

In the current analysis, the association between depression and cardiovascular events was strongest among individuals from areas with high deprivation and among individuals with low educational attainment. However, formal evidence for multiplicative and additive interaction was present for depression and area-based deprivation only. Existing publications on the effect of socioeconomic factors on the association between depression and cardiovascular disorders are conflicting. Whilst some studies have shown that individuals from a lower socioeconomic background might be more vulnerable to the adverse effect of depression (Avendano et al, 2006; Lazzarino et al, 2013a; b), one study did not observe any effect modification by socioeconomic factors (Mittag & Meyer, 2012), and one study showed that individuals from a high socioeconomic background are more vulnerable to the adverse effect of depression (Sun et al, 2016). Psychological distress was more strongly associated with all-cause mortality (Lazzarino et al, 2013a) and stroke and CHD mortality (Lazzarino et al, 2013b) among individuals with lower SES. The product term between depression and SES was significant at the 5% significance level in both publications. In keeping with that, Avendano et al (2006) showed that the hazard of stroke was higher among participants with depression and 10 – 12 years of school, 8 – 10 years of school, and 0 – 7 years of school (HR: 1.55, 95% CI: 0.79 – 3.04, and 1.46, 95% CI: 0.72 – 2.95, and 1.52, 95% CI: 0.73 – 3.15), relative to participants with at least 13 years of school. However, the CIs were wide and crossed the null value indicating that these results might be due to chance alone. Sun et al (2016) investigated to what extent the association between clinical depression and fatal or non-fatal stroke differed by education, income, and region in a longitudinal analysis using data from the China

Kadoorie Biobank. In contrast to the current analysis and the analyses by Lazzarino et al (2013a; 2013b) and Avendano et al (2006), Sun et al (2016) observed stronger associations between depression and stroke among individuals with high income, high educational attainment, and among participants from urban areas. Formal evidence of multiplicative interaction was present between depression and education ($p = 0.05$) but not depression and income ($p = 0.49$) or depression and area of residence ($p = 0.54$). Mittag & Meyer (2012) concluded that the association between depressive symptoms and IHD is not moderated by education. The author's conclusion was based on a non-significant interaction term (p-value not reported).

There are multiple possible explanations for the observed differences. First, the country of the study might have influenced the findings. Lazzarino et al (2013a; 2013b) findings are based on a study from England, Avendano et al (2006) and Mittag & Meyer (2012) used data from US participants, and Sun et al (2016) used data from Chinese participants. It is possible that results of studies conducted in high income countries such as the UK and US are not generalisable to findings in low income countries or findings in middle-income countries with periods of rapid growth such as China. Second, the assessment of exposure, outcome and socioeconomic factors differed across studies. This likely introduced heterogeneity and decreased comparability between studies. Furthermore, there was evidence of additive and multiplicative interaction between depression and area-based deprivation but neither evidence of additive nor multiplicative interaction between depression and educational attainment in the current analysis. This highlights that there might be differences in the effect of different socioeconomic factors on the association between depression and cardiovascular events. Third, the observed differences between groups might be due to chance alone. The CIs of the analyses were all wide indicating imprecise estimation. Moreover, some p-value did not reach the threshold for statistical significance indicating that there might not be any differences between groups, or that the study was not sufficiently powered to detect true differences between groups.

4.5.3.4 The role of comorbidities

There was evidence for effect modification by hypertension in the current analysis which is in contrast to existing studies. Sun et al (2016) investigated to what extent the association between clinical depression and first ever fatal or non-fatal stroke was modified by hypertension in the China Kadoorie Biobank cohort. The estimates in the subgroups with and without hypertension were very similar (HR: 1.13, 95% CI: 0.95 – 1.35, and 1.17, 95% CI: 0.90 – 1.53, respectively), and there was no evidence for multiplicative interaction between depression and hypertension (p = 0.83). Likewise, Pan et al (2011a) and Péquignot et al (2016) concluded that there was no effect modification by hypertension between depressive symptoms and stroke (p = 0.90) and depressive symptoms and combined CHD and stroke (p = 0.15), respectively. Again estimates of the association between depressive symptoms and stroke were very similar in participants with and without hypertension (HR: 1.29, 95% CI: 1.10 – 1.51, and 1.28, 95% CI: 0.96 – 1.69, respectively) in the analysis by Pan et al (2011a). Péquignot et al (2016) did not report estimates separately for individuals with and without hypertension at baseline. One potential explanation for the observed differences between the current analysis and the existing studies is that different scales were used to investigate interaction between depression and hypertension. Whilst the current analysis found evidence for additive interaction between depression and hypertension, there was no evidence for multiplicative interaction between depression and hypertension in analyses based on imputed data. The existing studies investigated to what extent there was evidence for multiplicative interaction by adding a product term to Cox proportional hazard models. However, they did not investigate to what extent there was evidence for additive interaction thereby potentially failing to identify that biological interaction between depression and hypertension was present (Andersson et al, 2005). Another potential reason for the observed differences between the current analysis and existing studies might be differences in determining hypertension. However, Sun et al (2016) assessed the presence of hypertension based on blood pressure measurements, self-reported diagnosis and treatment for hypertension which is in keeping with this analysis.

Similarly, Péquignot et al (2016) determined hypertension through blood pressure measurements and use of antihypertensive medication. Pan et al (2011a) determined hypertension solely based on a self-reported diagnosis of hypertension which might have been less accurate than the assessment methods of the other studies.

The results of existing studies on the association between depression and/ or diabetes and CVD are consistent with the current analysis. Scherrer et al (2011) investigated the hazard of MI among individuals with depression and/ or diabetes using Veterans Administration electronic medical records of 345,949 participants without CVD at baseline. In keeping with the current analysis, Scherrer et al (2011) found both depression alone and diabetes alone to be associated with increased hazard of MI relative to the group with neither condition, and the HR was highest among the group with comorbid depression and diabetes (1.82, 95% CI: 1.69 – 1.97). These authors did not calculate the excess risk due to interaction. I calculated the RERI of the estimates provided by Scherrer et al (2011) to be 0.2 but insufficient information was available to calculate CIs. Therefore, in contrast to this analysis there might have been a small supra-additive interaction. However, this might have been due to differences in the adjustment sets. Scherrer et al (2011) adjusted for age, sex, race, marital status, and insurance type whilst the current analysis additionally adjusted for comorbidities other than diabetes at baseline, as well as lifestyle factors such as smoking and physical activity and family history of CVD and/ or depression. Cummings et al (2016) investigated the effect of depression or stress on stroke, acute CHD and cardiovascular death separately for participants with and without diabetes among participants free from stroke or heart disease at baseline. Among participants with diabetes, elevated stress or depressive symptoms were associated with stroke and cardiovascular death and there was a non-statistically significant association with acute CHD in participants with diabetes in the fully adjusted model. Among those without diabetes, elevated stress or depressive symptoms were not associated with any of the outcomes, indicating that diabetes might be an effect modifier. The findings of the current analysis cannot be directly compared to the study by Cummings et al (2016) because they stratified their analysis by diabetes status, thereby creating a

comparison group that differed from the comparison group of the current analysis. However, the results indicated that there was an increased hazard of stroke among people with depression or stress and diabetes relative to the group with diabetes alone which is consistent with our findings. One study explored the individual and joint effect of antidepressant drugs and antidiabetic medication in a nationwide register study in Sweden (Rådholm et al, 2016). Although the exposure assessment differed from the current analysis, a combined treatment with antidepressant drugs and antidiabetic medication also showed a substantially increased hazard of first-ever fatal or non-fatal MI compared to individuals with treatment of neither or either condition alone. The corresponding HR for the combined effect of antidepressant drugs and antidiabetic medication compared to neither treatment for men and women aged 45 – 64 years were considerably higher than those of the current analysis (3.1, 95% CI 2.8 – 3.6, and 7.4, 95% CI: 6.3 – 8.6, respectively).

To the best of my knowledge, only one study investigated to what extent high cholesterol levels modified the association between depressive symptoms and fatal or non-fatal stroke. In keeping with the results of this study, Pan et al (2011a) did not find any evidence for multiplicative interaction between depressive symptoms and high cholesterol levels ($p = 0.93$). Furthermore, the stratified analysis showed very similar estimates for participants with and without high cholesterol levels (HR: 1.27, 95% CI: 1.08 – 1.48, and 1.36, 95% CI: 1.03 – 1.79, respectively), indicating little evidence for effect modification by high cholesterol levels.

4.6 Conclusion

This analysis made use of unique advantages of the UK Biobank, in particular its large sample size and the large amount of data collected. Depression, antidepressant use, hospital diagnosis with depression, and self-reported depression were associated with increased hazards of MCVE, even after adjustment for a wide range of potential confounding factors. Residual confounding due to measurement error of included covariates and due to unmeasured covariates might have introduced some bias. However, given the strength and consistency of the association, it was unlikely that

residual confounding alone explained the observed associations given the wide range of confounding factors considered. Furthermore, this analysis identified characteristics of individuals with depression which put them at particularly high hazard of MCVE. However, one limitation of this analysis was that, similar to what has been observed in the majority of existing studies, information on exposures and covariates was available at baseline only.

Chapter 5: Identification of depressive symptom trajectories and related cardiovascular risk factor profiles in young women

5.1 Background

In most existing studies on the association between clinical depression or depressive symptoms and subsequent cardiovascular events researchers relied on a measurement of the exposure at one point in time. Implicitly, these researchers assumed that a single measure of depression was a good approximation of the exposure status of participants throughout follow-up. However, this assumption may not be valid. For example, in a community-based sample Ames & Leadbeater (2018) identified four latent classes with different patterns of depressive symptoms among youth. Two subgroups were characterised by persistently high and low depressive symptoms (9% and 49%, respectively) but a considerable proportion of participants showed increasing and decreasing depressive symptoms (21% and 22%, respectively) from the age of 12 to 28 years. Whilst a single assessment of depressive symptoms would have accurately reflected the participant's exposure status with persistently high and low depressive symptoms, it would not accurately reflect the pattern of depressive symptoms of the latter two groups. As a result, researchers have emphasised the importance of considering the course of depressive symptoms over the life course (Colman & Ataullahjan, 2010; Musliner et al, 2016).

Unlike many other cohort studies, the ALSWH has assessed information on depressive symptoms, covariates and cardiovascular events at multiple occasions over several years. The ALSWH is an ongoing prospective cohort study of over 57,000 women in three age-cohorts that were recruited in 1996 and one additional younger cohort that was recruited later (Lee et al, 2005; Loxton et al, 2017). Women in the young, middle-age, and oldest cohort were aged 18 – 23 years, 45 – 50 years, and 70 – 75 years at baseline, respectively. Women were randomly selected from the national Medicare health insurance database, which includes all Australian citizens and permanent residents. Women living in rural and remote areas were intentionally

oversampled. Self-administered questionnaires were first sent to women in 1996, four years later, and every three years thereafter until 2015 (Lee et al, 2005). In 2012/ 2013 an additional age-cohort was recruited that consists of women born from 1989 to 1995. Due to its repeat assessments the ASLWH offered unique advantages with respect to assessing the patterns of depressive symptoms over time, exploring the risk of CVD that was associated with changes of depressive symptoms over time, and investigating the contribution of behavioural factors as potential mediating factors.

A previous study based on the mid-aged ALSWH cohort showed that although time-varying depressive symptoms were associated with increased risk of hypertension in the age-adjusted model, this association was no longer statistically significant in the fully adjusted model (Jackson et al, 2016). However, some of the factors adjusted for might have been on the causal pathway between depressive symptoms and hypertension. One of the candidate factors was BMI. Since many women were already overweight at baseline, the authors concluded that it was not appropriate to perform a mediation analysis between depressive symptoms, BMI and hypertension in this cohort. Therefore, it was of interest to investigate this further using data from women of the young cohort, where the temporality between depression and BMI could be more readily determined. Whilst using data from the young cohort of the ALSWH was advantageous with regard to the temporal relationship of variables, a disadvantage was that women of the young cohort were only aged 37 – 42 years at their last follow-up assessment. Thus, there were insufficient major CVD events to explore the relationship between repeated measures of depressive symptoms and CVD. As a result, I decided to use hypertension as outcome in this analysis because it develops at an earlier age and is a major risk factor for CVD events.

Using data from the young cohort of the ALSWH, the original plan at the start of the PhD was to identify subgroups of women with similar patterns of depressive symptoms over time, to perform a marginal structural model on the risk of hypertension associated with different patterns of change of depressive symptoms over time (Fewell et al, 2004; Gilsanz et al, 2015; Robins et al, 2000), and to investigate

the role of BMI as potential mediating factor in the relationship between depressive symptoms and hypertension. I intended to apply the counterfactual approach to causation to both the marginal structural model and the mediation analysis. I then explored this option in terms of feasibility and methodology in the context of psychological disorders and decided not to pursue these analyses within the PhD (see section 7.3.2 The need for an accurate definition of depression for more detailed discussion). As a result, this analysis focusses on identifying women with similar depressive symptoms trajectories over time and on exploring their profiles of key cardiovascular risk factors at the beginning and end of follow-up.

Group-based trajectory modelling aims to identify subgroups of individuals with similar patterns of an exposure over time by simplifying heterogeneous populations into more homogenous subgroups (Lennon et al, 2018). Group-based trajectory modelling has multiple advantages over modelling patterns of change among groups that have been specified *a priori*. For example, since no assumption about the number of groups or shapes of trajectories are made prior to the analysis, it allows for a one-class solution in which all individuals of the sample follow similar patterns of depressive symptoms over time. Another advantage in the presence of at least two subgroups is that it allows for an investigation of the certainty with which participants could be allocated to one of the identified classes (Nagin & Odgers, 2010).

The identification of subgroups with similar trajectories over time might be particularly valuable in the context of depression (Schubert et al, 2017). Due to the substantial heterogeneity between participants with depressive symptoms, different patterns of depressive symptoms over time might be indicative of different causes and consequences. To investigate whether the identified subgroups were clinically meaningful, I investigated how the identified subgroups were related to key cardiovascular risk factors at the beginning and end of follow-up. Due to the potentially important role of BMI as mediating factor, I additionally investigated patterns of change of BMI over time for each of the identified subgroups and explored differences across groups.

5.2 Objectives

1. To identify subgroups of women with similar patterns of depressive symptoms over time using group-based trajectory modelling
2. To investigate whether these subgroups had
   a) different profiles of key cardiovascular risk factors at the beginning and end of follow-up
   b) different patterns of change of BMI over time

5.3 Methods

5.3.1 Sample

This analysis included women of the young cohort born from 1973 to 1978, who were aged 18 – 23 years at the start of the study and 37 – 42 years at the end of follow-up. Since depressive symptoms were first assessed in the second study phase, women were excluded from the analysis if they only participated in phase 1. Participants with no information on depressive symptoms in any of the study phases were excluded from the analyses. Women who had not responded to survey 2 were contacted and offered to complete a shorter version of the questionnaire that contained questions that were considered most important. Following recommendations from the ALSWH study team, participants who only completed the short survey at study phase 2 were excluded from this analysis. Observations of women were included until the end of follow-up, time of death, or date of withdrawal. Observations for which the date of withdrawal was the same as the date on which the survey was returned were included in the analysis.

Secondary analyses of ALSWH data were conducted under generic approval from the University of Newcastle's Human Research Ethics Committee, approval numbers H-076-0795 and H-2012-0256, and the University of Queensland's Medical Research Ethics Committee, approval numbers 2004000224 and 2012000950 (all approved in 1995). All participants gave informed consent to be included in the study.

5.3.2 Variables

Depressive symptoms were first assessed in the second study phase and in every survey thereafter using the 10-item version of the CES-D (Andresen et al, 1994; Radloff, 1977). Women were asked to rate the frequency of different depressive symptoms in the past week on a scale of: '0 = rarely or none of the time (less than 1 day)'; '1 = some or a little of the time (1 – 2 days)'; '2 = occasionally or a moderate amount of the time (3 – 4 days)'; or '3 = most or all of the time (5 – 7 days)'. In keeping with recommendations by Andresen et al (1994), information on depressive symptoms was set to missing if fewer than nine CES-D items were completed. If women provided information on nine items, the value of the tenth item was considered to be the mean of the nine items provided. As a result, the final score ranges from 0 to 30 with higher scores indicating greater depressive symptoms. The CES-D score was treated as continuous variable when using group-based trajectory modelling. In addition, a binary variable was created to indicate the presence or absence of depressive symptoms in descriptive statistics. In keeping with recommendations (Andresen et al, 1994) and a previous ALSWH publication (Jackson et al, 2016), a score of at least 10 was classified as presence of depressive symptoms.

Sociodemographic factors included age, area of residence, living arrangements, educational attainment, and ability to manage on income. Area of residence was categorised as metropolitan, rural and remote. Marital status was grouped into three categories of married/ de-facto, separated/ divorced/ widowed, and single. Women were asked who lived with them. That information was used to create a binary variable indicating whether or not women lived alone. SES was determined by two measures. First, women reported their highest educational attainment. Due to the large proportion of women reporting a university degree or higher university degree, educational attainment was treated as binary variable indicating whether or not women had a (higher) university degree. Second, women were asked how well they managed on their income. This information was grouped into three categories of easy/ not too bad, difficult some of the time and difficult all of the time/ impossible.

Since women were not asked how well they managed on their income in study phase 2, information provided in study phase 1 was carried forward to the next phase.

Health behaviours included smoking, physical activity, and alcohol intake. Smoking was classified as never, ex-smoker, or current smoker. Physical activity was assessed using the metabolic equivalent (MET), a measure to define levels of physical activity as multiples of resting metabolic rate. Women were asked to report the time spent walking, in moderate leisure activity and vigorous leisure activity in the last week. A MET value of 3.33 has been assigned for walking and moderate leisure activity, a value of 6.66 has been assigned to vigorous physical activity. MET minutes per week were calculated as 3.33 * minutes walking + 3.33 * minutes moderate activities + 6.66 * minutes vigorous activities (Brown & Pavey, 2016). Using that information women's physical activity was defined as nil/ sedentary (0 to < 33 MET minutes/ week), low (33.3 to < 500 MET minutes/ week), moderate (500 to < 1000 MET minutes/ week), and high (≥ 1000 MET minutes/ week). Alcohol intake was defined in keeping with the Australian National Health and Medical Research Council guidelines (National Health and Medical Research Council, 2001). Due to the small proportion of risky drinkers (15 – 28 drinks/ week) and high-risk drinkers (> 28 drinks/ week), these groups were combined. Furthermore, women who reported that they drink less than once a month and non-drinkers were classified as non/ rarely drinkers. The remaining women were classified as low-risk drinkers (up to 14 drinks/ week).

Variables related to the health status of participants included BMI, hypertension, diabetes, or heart disease. BMI (kg/m^2) was calculated by the ALSWH study team based on women's self-reported weight and height. The ALSWH study team set the BMI of pregnant women to missing from study phase 1 to 3 whereas it was set to the pre-pregnancy BMI from study phase 4 onwards. BMI was defined as continuous and categorical variable in keeping with WHO recommendations (underweight: < 18.5 kg/m^2, normal weight: 18.5 – 24.9 kg/m^2, overweight: 25 – 29.9 kg/m^2 or obese: ≥ 30 kg/m^2) (World Health Organisation, 2000). Women were asked to report a diagnosis with or treatment for hypertension, diabetes or heart disease since the previous

survey. Once women first reported occurrence of either of these medical conditions, they were considered to have it at each subsequent survey.

5.3.3 Statistical analyses

5.3.3.1 Descriptive statistics

Participant characteristics were described at each study phase and characteristics of included and excluded participants were compared. Since a large proportion of excluded women only participated in phase 1, the comparison of those included versus excluded was made with regard to participant characteristics at study phase 1. Differences between those included versus excluded were assessed using t-tests for normally distributed continuous variables and chi-squared tests for proportions. A two-sided p-value < 0.05 was considered statistically significant. Using the binary variable for absence/ presence of depressive symptoms, patterns of depressive symptoms from study phases 2 to 7 were investigated. The number of unique response patterns was determined and the most common response patterns were reported. A random sample of 250 women was selected to illustrate individual depressive symptom trajectories plotted as a function of age using the continuous depressive symptom score.

5.3.3.2 Group-based trajectory modelling of depressive symptoms (CES-D)

5.3.3.2.1 Latent process mixed modelling

Subgroups of participants with similar depressive symptoms trajectories were identified using the lcmm function of the R package Latent Class Mixed Modelling (LCMM) (Proust-Lima et al, 2017). Latent class mixed modelling assumes that the study population is heterogeneous and formed of a specific number of latent classes whose mean trajectory profiles can be estimated. The groups are referred to as latent classes because they are neither directly observed nor defined *a priori*. The heterogeneity between trajectories of participants is taken into account and subgroups with similar trajectories are identified. Latent process mixed models are an extension of standard linear mixed models that allow the study of non-Gaussian longitudinal markers over time. To estimate latent process linear mixed models, one needs to specify fixed effects (linear regression at the population level), mixture

effects (subset of covariates having a class-specific effect), and random effects (subset of covariates having a subject-specific effect in order to allow individual variation within classes and to take account of correlation between repeated measures of the same individual). Missing data are handled by fitting the model using maximum likelihood estimation (Nagin & Odgers, 2010). Subjects with missing data are included in the analysis but only available data for each subject are used. This generates unbiased parameter estimates under the assumption that data are MAR (Nagin & Odgers, 2010). Nagin (2005) highlighted that a sample of at least 300 to 500 subjects is required to estimate trajectory models. Bigger samples are preferable to increase statistical power and ease identification of trajectory groups. The minimal number of data points depends on the complexity of the model. Whilst three data points might be sufficient to estimate linear and quadratic trajectories, preferably the dataset should have at least four or five data points to estimate more complex models (Curran & Muthén, 1999).

5.3.3.2.2 Link function

The optimal model to identify subgroups of participants with similar depressive symptom trajectories was chosen following a framework to construct latent class trajectory models developed by Lennon et al (2018); however, the first step of the framework was altered to take account of specific properties of the depressive symptom scale. As described previously, the depressive symptom score used in this analysis is a sum score of ten items. Proust-Lima et al (2011) emphasised that these discrete quantitative outcomes often suffer from floor and/ or ceiling effects due to the limited range of possible values, and they suffer from curvilinearity due to varying sensitivity to change at different levels of the scale. As a result, the relationship between depressive symptom scores and the underlying latent process of interest might not be accurately described through a linear link function (which is the default option in R software). Thus, the following method was developed based on recommendations by Proust-Lima et al (2011; 2017). Five latent process mixed models with depressive symptom scores as dependent variable were fitted using different link functions (linear, rescaled beta cumulative distribution function, five

equidistant I-splines, five I-splines set at the quantiles of the distribution (Ramsay, 1988), thresholds). The link function that provided the best model fit was determined using the discrete Akaike information criterion (AIC). In contrast to the AIC and Bayesian information criterion (BIC) for continuous outcomes, the discrete AIC takes into account that the measure of depressive symptoms is a discrete and bounded quantitative outcome by using the discrete log-likelihood instead of the continuous log-likelihood to estimate the information criterion. The fixed and random effects structures were kept constant in these models. Time was modelled using mean-centred age divided by ten in order to ease the interpretation of the intercepts and reduce numerical problems due to very large numbers in quadratic models. Natural cubic splines of age were used to model time dependence for the fixed effects with the spline knots set at the quartiles of the distribution of age. A random intercept and slope was added to account for inter-subject variability. Due to the computational complexity of the models, the best link function was determined in models with one class. The model with the preferred link function was treated as scoping model. The R code for the model using a linear link function is presented below as an example.

```
kn <- quantile(sample$age, probs = c(0.25,0.50,0.75))      #spline knots at the quartiles
                                                           #of the distribution

lcmm(fixed = CESD ~ ns(age, knots = kn),
random = ~ time,
subject = "ID",
link = "linear",          #five different models using different link functions were fitted:
                          #linear, beta, equidistant I-splines, I-splines set at quantiles
                          #of distribution, threshold
data = sample)
```

5.3.3.2.3 Number of classes

After selecting the preferred link function, the optimal number of classes was identified. In this step, models with increasing numbers of groups were fitted whilst the link function and structure of the fixed and random effects were kept the same. To model time dependence for the mixture effects in models with more than one class natural cubic splines of age were used as before with spline knots set at the quartiles of the distribution. In the presence of at least two groups and mixture effects, it is recommended that the same model should be fitted several times using different sets of starting values. This is because the log-likelihood might have multiple maxima and convergence towards the global maximum is not ensured when running the model with only one set of starting values (Proust-Lima et al, 2017). Consequently, all models with at least two classes were run several times using different sets of starting values that were chosen at random from values of the asymptotic distribution of the maximum likelihood estimates of the one class model with the same model structure. Models with increasing number of classes were fitted until two models failed to reach convergence in 500 iterations. As recommended by van de Schoot et al (2017), the optimal number of classes to describe the dataset was determined by balancing statistical criteria (lowest discrete AIC) against criteria reflecting the clinical relevance and usefulness of the identified subgroups (e.g. group size).

5.3.3.2.4 Model structure

After selecting the optimal number of groups, the model structure was refined. The optimal model structure was determined by fitting models with increasingly complex model structures (Table 5.1). The link function and number of classes were kept constant in these models. The optimal model structure was determined based on the lowest discrete AIC of all models that reached convergence in less than 500 iterations.

Table 5.1: Models with increasingly complex model structures

Model description	Fixed effects	Mixture effects	Random effects	Class-specific variance-covariance matrix
Fixed effects	Natural cubic splines of age	Natural cubic splines of age	---	No
Random intercept	Natural cubic splines of age	Natural cubic splines of age	Intercept	No
Random slope	Natural cubic splines of age	Natural cubic splines of age	Intercept & slope	No
Random quadratic - common variance structure	Natural cubic splines of age	Natural cubic splines of age	Intercept, slope & quadratic	No
Random quadratic - variance structure varies across classes with proportionality constraint	Natural cubic splines of age	Natural cubic splines of age	Intercept, slope & quadratic	Yes (proportionality constraint)

5.3.3.2.5 Model adequacy

The favoured model was chosen based on the results of the assessment of the optimal structure and results of model adequacy tools. Model adequacy was investigated for each model of Table 5.1 that reached convergence in less than 500 iterations following guidance in the literature (Lennon et al, 2018). First, the number and proportion of women in each of the classes was described. In keeping with recommendations by Lennon et al (2018), models were regarded acceptable if at least 1% of the sample was allocated to each of the classes to ensure clinical plausibility and meaningfulness of the groups. Second, the average of the posterior probability of assignment (APPA) was estimated for each of the classes. For this measure the posterior probability of belonging to each of the classes was estimated for all women. Women were allocated to the class with the highest individual posterior class membership probability. The APPA is the average of all posterior class membership probabilities of women allocated to that class (Lennon et al, 2018). In keeping with Lennon et al (2018), a threshold of above 70% for all classes was regarded as acceptable. If this threshold was not reached, the average posterior probabilities of belonging to any of the other classes were investigated. Third, relative entropy was estimated as a measure of classification uncertainty. Relative entropy can take values between 0 and 1 with

higher values indicating less classification uncertainty. In keeping with recommendations by Lennon et al (2018), relative entropy values greater than 0.5 were considered acceptable.

5.3.3.2.6 Clinical characterisation and plausibility

The clinical characterisation and plausibility of the trajectory groups was assessed. First, the mean predicted trajectories of depressive symptoms over time and 95% prediction intervals were plotted for each identified class of the favoured model. Second, spaghetti plots of individual depressive symptom trajectories of 100 randomly selected women allocated to each of the classes were plotted. Third, the extent to which the trajectory groups are distinguishable in terms of pre-existing characteristics and characteristics at the end of follow-up. For this purpose, participant characteristics at baseline (study phase 2) and end of follow-up (study phase 7) were investigated separately for each of the trajectory groups among women that had measures at both time points. To test for between-group differences chi-squared tests with continuity correction were used for categorical variables, and t-tests were used for normally distributed continuous variables (Yoshida et al, 2019). For non-normally distributed continuous variables, Kruskal-Wallis tests were used (Yoshida et al, 2019).

5.3.3.3 Class-specific trajectories of BMI

Among participants with at least one measure of BMI, the class-specific mean predicted trajectory of BMI was modelled. Using the Wald test, I investigated whether class membership was associated with different patterns of change of BMI over time. A two-sided p-value of <0.05 was considered statistically significant.

5.3.3.4 Sensitivity analysis

First, since some existing studies have excluded participants with less than two assessments of depressive symptoms throughout follow-up (Leigh et al, 2016; Nabi et al, 2008; Tran et al, 2019), the analysis was repeated after exclusion of 1,352 women with only one assessment of depressive symptoms. Second, to meaningfully interpret changes in depressive symptoms over time, two assumptions have to be met: the

depressive symptom scale has to measure a single construct (unidimensionality) and has to measure that construct the same way across time (measurement invariance) (Fried et al, 2016; van de Schoot et al, 2017). It has been shown that these assumptions might not hold for depressive symptom rating scales (Fried et al, 2016). To overcome this problem, investigating specific depressive symptoms instead of investigating a sum score of multiple depressive symptoms has been recommended (Fried & Nesse, 2015b). Therefore, the analysis was repeated using individual items of the CES-D scale. Due to time restrictions I decided to choose two out of the ten CES-D items for this analysis, as follows: "I felt depressed" and "I felt hopeful about the future". These items were chosen because they have been shown to be among the specific symptoms of depression most commonly captured by depressive symptom scales (Fried, 2017a). "I felt hopeful about the future" is a positive mood item on the CES-D which was reversed for the analysis.

5.4 Results

5.4.1 Study population

Out of 14,247 participants, 11,804 women were included in the analysis. Reasons for exclusion were participation only in study phase 1, withdrawal date earlier than date of participation, no depressive symptom measure in any of the study phases, and completion of the short survey at study phase 2 (Figure 5.1).

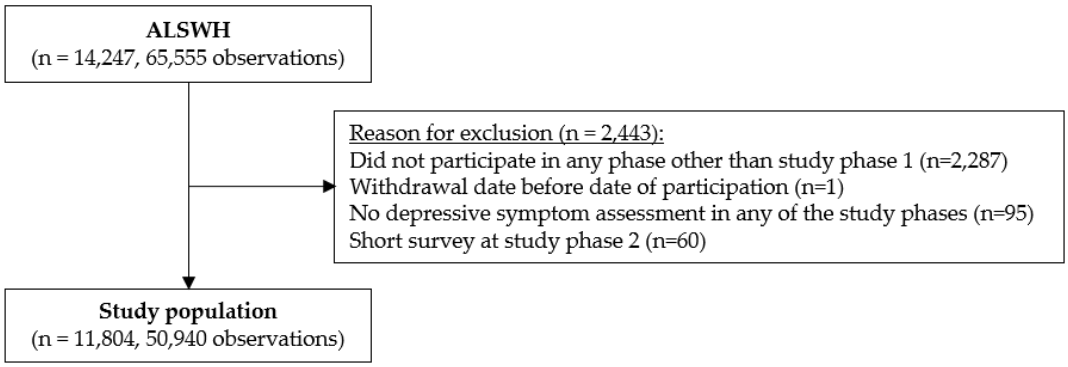


Figure 5.1: Flow chart of Australian Longitudinal Study on Women's Health participants included in analyses

There were statistically significant differences between included and excluded women in all variables except for living arrangements, area of residence, and history

of heart disease (Appendix Table A 35). The proportion of women with a university degree was lower and the proportion of women who reported that it was difficult to manage on their income was higher among women excluded from the analysis. Furthermore, those excluded were more likely to report that they were married or separated/ divorced/ widowed, current smokers, and to have risky alcohol intake, hypertension or diabetes than participants who were included in the analysis. The mean BMI was slightly lower among those excluded than among those included in the analysis; however, this might have been driven by the higher proportion of women with underweight among those excluded from the analysis. Although there were statistically significant age differences, the difference in the median was small.

5.4.2 Descriptive statistics

Out of 11,804 eligible women, 9,560 (81.0%) participated at study phase 2 (hereafter referred to as baseline) and 7,149 (60.6%) participated at study phase 7 (hereafter referred to as end of follow-up). At baseline, the median (IQR) age of women was 24.6 (23.3, 25.8) years and 45.0% were married. Over half of the women lived in metropolitan areas (54.7%), 39.0% had a university degree, and 50.9% reported it was easy/ not too bad to manage on their income. The proportions of women with a university degree; who were married; found it easy/ not too bad to manage on their income; had risky/ high risk alcohol intake; sedentary physical activity levels; and who were overweight or obese increased over time. In contrast, the proportion of current smokers decreased over time. The proportion of women with diabetes and heart disease remained low throughout follow-up (3.0% and 1.0% at the end of follow-up, respectively). Whilst 251 (2.6%) women reported hypertension at baseline, 710 (9.9%) women reported hypertension at the end of follow-up.

*Table 5.2: Participant characteristics of the ALSWH cohort at each study phase (page 1 of 2)**

	Phase 2 (n=9,560)	**Phase 3** (n=9,022)	**Phase 4** (n=9,092)	**Phase 5** (n=8,152)	**Phase 6** (n=7,965)	**Phase 7** (n=7,149)
Age (median [IQR])	24.6 [23.3, 25.8]	27.6 [26.3, 28.8]	30.6 [29.3, 31.8]	33.7 [32.5, 34.9]	36.7 [35.5, 38.0]	39.6 [38.4, 40.9]
Marital status (%)						
Married/ de-facto	4,301 (45.0)	5,519 (61.2)	6,550 (72.0)	6,283 (77.1)	6,195 (77.8)	5,406 (75.6)
Separated/ divorced/ widowed	144 (1.5)	334 (3.7)	408 (4.5)	446 (5.5)	577 (7.2)	568 (7.9)
Single	5,071 (53.0)	3,133 (34.7)	2,095 (23.0)	1,394 (17.1)	1,098 (13.8)	813 (11.4)
Missing value	44 (0.5)	36 (0.4)	39 (0.4)	29 (0.4)	95 (1.2)	362 (5.1)
Lives alone (%)						
No	8,774 (91.8)	8,310 (92.1)	8,304 (91.3)	7,463 (91.5)	7,273 (91.3)	6,326 (88.5)
Yes	586 (6.1)	678 (7.5)	767 (8.4)	678 (8.3)	605 (7.6)	462 (6.5)
Missing value	200 (2.1)	34 (0.4)	21 (0.2)	11 (0.1)	87 (1.1)	361 (5.0)
Area of residence (%)						
Metropolitan	5,225 (54.7)	5,200 (57.6)	5,380 (59.2)	4,756 (58.3)	4,622 (58.0)	4,053 (56.7)
Rural	3,931 (41.1)	3,484 (38.6)	3,215 (35.4)	2,905 (35.6)	2,847 (35.7)	2,418 (33.8)
Remote	363 (3.8)	335 (3.7)	381 (4.2)	303 (3.7)	278 (3.5)	197 (2.8)
Missing value	41 (0.4)	3 (0.0)	116 (1.3)	188 (2.3)	218 (2.7)	481 (6.7)
University degree (%)						
No	5,484 (57.4)	4,906 (54.4)	4,791 (52.7)	3,834 (47.0)	3,579 (44.9)	2,899 (40.6)
Yes	3,733 (39.0)	3,906 (43.3)	4,268 (46.9)	4,144 (50.8)	4,243 (53.3)	3,870 (54.1)
Missing value	343 (3.6)	210 (2.3)	33 (0.4)	174 (2.1)	143 (1.8)	380 (5.3)
Ability to manage on income (%)						
Difficult/ impossible	1,603 (16.8)	1,074 (11.9)	1,141 (12.5)	992 (12.2)	1,038 (13.0)	982 (13.7)
Difficult some of the time	3,063 (32.0)	2,706 (30.0)	2,630 (28.9)	2,282 (28.0)	2,426 (30.5)	1,880 (26.3)
Easy/ not too bad	4,864 (50.9)	5,196 (57.6)	5,280 (58.1)	4,843 (59.4)	4,405 (55.3)	3,924 (54.9)
Missing value	30 (0.3)	46 (0.5)	41 (0.5)	35 (0.4)	96 (1.2)	363 (5.1)
Alcohol intake (%)						
Low risk	5,542 (58.0)	5,491 (60.9)	5,440 (59.8)	4,810 (59.0)	4,596 (57.7)	4,061 (56.8)
Non/ rarely	3,597 (37.6)	3,168 (35.1)	3,267 (35.9)	2,948 (36.2)	2,935 (36.8)	2,405 (33.6)
Risky/ high risk	359 (3.8)	327 (3.6)	340 (3.7)	361 (4.4)	363 (4.6)	464 (6.5)
Missing value	62 (0.6)	36 (0.4)	45 (0.5)	33 (0.4)	71 (0.9)	219 (3.1)
Smoking status (%)						
Never-smoker	5,443 (56.9)	5,139 (57.0)	5,254 (57.8)	4,852 (59.5)	4,828 (60.6)	4,270 (59.7)
Ex-smoker	1,376 (14.4)	1,664 (18.4)	2,008 (22.1)	2,092 (25.7)	2,157 (27.1)	1,932 (27.0)
Current smoker	2,664 (27.9)	2,187 (24.2)	1,784 (19.6)	1,190 (14.6)	923 (11.6)	720 (10.1)
Missing value	77 (0.8)	32 (0.4)	46 (0.5)	18 (0.2)	57 (0.7)	227 (3.2)
Physical activity (%)						
Nil/ sedentary	934 (9.8)	815 (9.0)	1,022 (11.2)	1,121 (13.8)	1,202 (15.1)	1,079 (15.1)
Low	3,057 (32.0)	2,919 (32.4)	3,130 (34.4)	2,913 (35.7)	2,687 (33.7)	1,911 (26.7)
Moderate	2,221 (23.2)	2,104 (23.3)	2,103 (23.1)	1,777 (21.8)	1,698 (21.3)	1,447 (20.2)
High	3,305 (34.6)	3,118 (34.6)	2,707 (29.8)	2,217 (27.2)	2,165 (27.2)	2,176 (30.4)
Missing value	43 (0.4)	66 (0.7)	130 (1.4)	124 (1.5)	213 (2.7)	536 (7.5)

*Table 5.2: Participant characteristics of the ALSWH cohort at each study phase (page 2 of 2)**

	Phase 2 (n=9,560)	**Phase 3** (n=9,022)	**Phase 4** (n=9,092)	**Phase 5** (n=8,152)	**Phase 6** (n=7,965)	**Phase 7** (n=7,149)
Body mass index (kg/m^2) (mean (SD))	24.0 (5.0)	24.8 (5.5)	25.4 (5.9)	26.0 (6.0)	26.4 (6.3)	27.0 (6.6)
Body mass index (WHO groups) (%)						
Underweight	569 (6.0)	366 (4.1)	312 (3.4)	210 (2.6)	191 (2.4)	123 (1.7)
Acceptable weight	5,465 (57.2)	4,673 (51.8)	4,840 (53.2)	4,121 (50.6)	3,795 (47.6)	3,094 (43.3)
Overweight	1,743 (18.2)	1,791 (19.9)	2,143 (23.6)	2,051 (25.2)	2,025 (25.4)	1,875 (26.2)
Obese	969 (10.1)	1,223 (13.6)	1,581 (17.4)	1,629 (20.0)	1,798 (22.6)	1,790 (25.0)
Missing value	814 (8.5)	969 (10.7)	216 (2.4)	141 (1.7)	156 (2.0)	267 (3.7)
Depressive symptom score (median [IQR])	6.0 [3.0, 10.0]	6.0 [3.0, 10.0]	5.0 [3.0, 9.0]	5.0 [3.0, 9.0]	5.0 [2.0, 8.0]	5.0 [2.0, 9.0]
Depressive symptoms (binary) (%)						
No	6,526 (68.3)	6,565 (72.8)	6,904 (75.9)	6,253 (76.7)	6,251 (78.5)	5,191 (72.6)
Yes	2,803 (29.3)	2,311 (25.6)	2,064 (22.7)	1,773 (21.7)	1,594 (20.0)	1,652 (23.1)
Missing value	231 (2.4)	146 (1.6)	124 (1.4)	126 (1.5)	120 (1.5)	306 (4.3)
Diabetes (%)						
No	9,429 (98.6)	8,914 (98.8)	8,924 (98.2)	7,968 (97.7)	7,754 (97.4)	6,935 (97.0)
Yes	29 (0.3)	77 (0.9)	117 (1.3)	160 (2.0)	206 (2.6)	214 (3.0)
Missing value	102 (1.1)	31 (0.3)	51 (0.6)	24 (0.3)	5 (0.1)	0 (0.0)
Heart disease (%)						
No	9,444 (98.8)	8,962 (99.3)	8,996 (98.9)	8,087 (99.2)	7,895 (99.1)	7,080 (99.0)
Yes	14 (0.1)	29 (0.3)	45 (0.5)	41 (0.5)	65 (0.8)	69 (1.0)
Missing value	102 (1.1)	31 (0.3)	51 (0.6)	24 (0.3)	5 (0.1)	0 (0.0)
Hypertension (%)						
No	9,207 (96.3)	8,658 (96.0)	8,638 (95.0)	7,566 (92.8)	7,289 (91.5)	6,439 (90.1)
Yes	251 (2.6)	333 (3.7)	403 (4.4)	562 (6.9)	671 (8.4)	710 (9.9)
Missing value	102 (1.1)	31 (0.3)	51 (0.6)	24 (0.3)	5 (0.1)	0 (0.0)

* Data are given as n (%) unless specified

Using the binary depressive symptoms variable, the proportion of women with depressive symptoms was 29.3% at baseline and 23.1% at the end of follow-up (Table 5.2). There were 666 unique depressive symptom response patterns over time, with the 20 most common patterns presented in Table 5.3. The most common pattern was that women did not report depressive symptoms at any study phase (1,892 women, 16.0%). The 18th most common pattern was that women reported depressive symptoms at all study phases (109 women, 0.9%).

Table 5.3: The 20 most common depressive symptom response patterns using the binary depressive symptom variable

Ranking	2nd wave	3rd wave	4th wave	5th wave	6th wave	7th wave	Count
1	No	No	No	No	No	No	1892
2	No	Missing	Missing	Missing	Missing	Missing	460
3	No	No	No	No	No	Missing	291
4	Yes	Missing	Missing	Missing	Missing	Missing	282
5	Yes	No	No	No	No	No	256
6	Missing	No	No	No	No	No	252
7	No	No	No	Missing	Missing	Missing	248
8	No	No	Missing	Missing	Missing	Missing	211
9	No	No	No	No	No	Yes	165
10	No	Yes	No	No	No	No	165
11	Missing	No	Missing	Missing	Missing	Missing	161
12	No	No	No	No	Missing	Missing	158
13	No	Missing	No	No	No	No	148
14	Missing	Missing	No	Missing	Missing	Missing	128
15	No	No	No	Missing	No	No	125
16	No	No	No	Yes	No	No	120
17	No	No	Yes	No	No	No	109
18	Yes	Yes	Yes	Yes	Yes	Yes	109
19	No	Missing	No	Missing	Missing	Missing	99
20	Yes	Yes	Missing	Missing	Missing	Missing	93

There was variation between depressive symptom scores reported by the same woman over time as well as variation between women (Figure 5.2). Throughout follow-up the majority of women reported depressive symptom scores below ten. There were more depressive symptom scores above ten among women aged 20 to 30 years than among women over 30 years. However, it should be noted that some of these higher scores were reported by women with one only measure of depressive symptoms which might indicate that women with higher depressive symptom scores at younger ages were lost to follow-up.

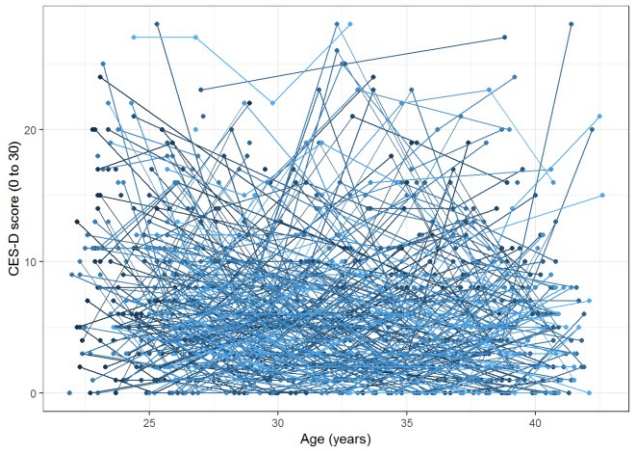


Figure 5.2: Depressive symptom trajectories of 250 randomly selected women as a function of age

5.4.3 Identifying subgroups of women with similar depressive symptom trajectories

5.4.3.1 Link function

In the first step, five models with different link functions were fitted to identify the link function that provided the best model fit. The approximation of the linear link function was relatively far from the approximation of the non-linear link functions. In contrast, the approximations of the non-linear link functions were all close indicating that they might be good alternatives to each other (Figure 5.3). The model with a thresholds link function had the lowest discrete AIC (Table 5.4). The difference in discrete AIC values (Δ AIC) was greatest between the model using the thresholds link function and the linear link function (Δ AIC: 12,014), and smallest between the model using the thresholds link function and the I-splines link function with five equidistant knots (Δ AIC: 251). Proust-Lima et al (2011) recommend choosing the optimal link function based on the lowest discrete AIC but the estimation of the model using the thresholds link function was computationally intensive. For example, the computation of a two-class model took more than 24 hours. Since the difference between the approximations of all non-linear models were small, I decided to select the I-splines link function with five equidistant knots as the preferred link function.

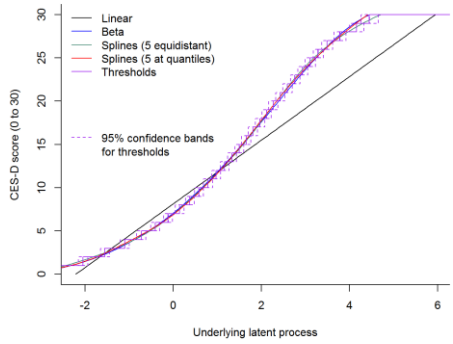


Figure 5.3: Estimated link functions of depressive symptoms (CES-D, range 0 to 30) in latent process mixed models

Table 5.4: Model fit of five latent process mixed models using different link functions

Link function	Number of iterations	Number of parameters	Discrete AIC	Δ AIC*
Linear	13	9	286,272	12,014
Beta	9	11	274,607	349
Splines (5 equidistant)	26	14	274,509	251
Splines (5 at quantiles)	17	14	274,553	296
Thresholds	15	37	274,257	---

* Difference in AIC between model using thresholds link function and models using other link functions

5.4.3.2 Number of groups

In the second step, the optimal number of groups was determined with the link function selected as I-splines with five equidistant knots. The lowest discrete AIC was observed in a model with five classes (Table 5.5); however, three of these classes contained less than 1% of the sample and the predicted mean trajectories of these groups largely overlapped (Appendix Figure A 1). The second lowest discrete AIC value was observed in a model with four groups. Again, two of the classes contained less than 1% of the sample and the predicted mean trajectories of the groups largely overlapped. The model with three groups had the third lowest discrete AIC value and lowest BIC value and whilst one of the groups was small, based on existing literature the small proportion of women with fluctuating depressive symptoms over time was considered plausible and potentially clinically relevant. Balancing the results of criteria reflecting clinical relevance/ usefulness and statistical criteria, the optimal number of classes was determined as three.

Table 5.5: Model characteristics of models with increasing numbers of latent classes

Number of groups	Number of iterations	Log-likelihood	Number of parameters	Discrete AIC	BIC	% class 1	% class 2	% class 3	% class 4	% class 5
1	26	-138,030	14	274,509	276,191	100.0				
2	36	-137,972	20	274,437	276,132	27.4	72.6			
3	330	-137,940	26	274,392	276,125	71.1	28.1	0.7		
4	220	-137,891	32	274,312	276,083	0.9	0.7	70.9	27.6	
5	214	-137,859	38	274,263	276,075	0.2	70.3	0.7	0.9	27.9
6	Model did not reach convergence in 500 iterations									
7	Model did not reach convergence in 500 iterations									

5.4.3.3 Model structure

With the link function selected as I-splines with five equidistant knots and the number of groups selected as three, five models with increasingly complex model structures were fitted (Table 5.6). Four of the five models reached convergence in less than 500 iterations. The model with quadratic random effects that allowed the variance-covariance structure of random effects to vary across classes with a proportionality constraint failed to reach convergence. The model with random quadratic effects and a common variance structure of random effects across classed yielded the lowest discrete AIC; however, the discrete AIC values of all random effect models were close to each other. The model with no random effect was the model with the highest discrete AIC.

Table 5.6: Model fit and model adequacy of models with increasingly complex model structures

Model structure	Discrete AIC	BIC	Proportion per class (%)	Average posterior probability of assignment per class (%)	Relative entropy
Fixed effect	275,768	277,317	51:22:28	80:84:86	0.64
Random intercept	274,455	276,178	5:22:73	61:65:74	0.39
Random slope	274,392	276,125	71:28:1	80:71:59	0.52
Random quadratic (common var-cov* matrix across classes)	274,369	276,109	1:30:69	62:72:80	0.52
Random quadratic (var-cov* matrix with proportionality constraint)	Model did not reach convergence in 500 iterations				

*var-cov: variance- covariance

5.4.3.4 Model adequacy

Model adequacy was assessed for all models of Table 5.6 that reached convergence. The proportion of women allocated to each of the classes was at least 1% in all of the models. The largest groups were made up of about 70% of all women in the random intercept, random slope and random quadratic model with a common variance-covariance structure across classes and of about 50% in the fixed effects model. The average posterior probability of assignment per class is recommended to be above 70% for each of the classes. This recommendation was met for the fixed effect model. However, it was not met for one class of the random slope and random quadratic

models nor was it met for two classes of the random intercept model. Using the model with the lowest discrete AIC as an example, the posterior classification table of the random quadratic model indicated that women allocated to the first class of the model had an average posterior probability of 62% of belonging to that class, an average posterior probability of 17% of belonging to the second class, and an average posterior probability of 20% of belonging to the third class. Women allocated to the second class of the model had an average posterior probability of 72% of belonging to that class, an average posterior probability of 3% of belonging to the first class, and an average posterior probability of 25% of belonging to the third class. Women allocated to the third class had an average posterior probability of 80% of belonging to that class, an average posterior probability of 3% of belonging to the first class, and an average posterior probability of 18% of belonging to the second class. The relative entropy values of the models ranged from 0.39 to 0.64, with the fixed effects model reaching the highest relative entropy value. Whilst the fixed effects, random slope, and random quadratic model with a common variance-covariance structure across classes met the recommended relative entropy value of at least 0.5, the relative entropy value of the random intercept model was below that threshold.

The favoured model was chosen balancing results of model adequacy tools and the assessment of the optimal structure. The fixed effect model was the only model that met all criteria of the model adequacy tools but was also the model with the highest discrete AIC value. The random quadratic model with a common variance-covariance structure across classes provided the lowest discrete AIC value; however, one of the classes had an APPA value below 70%. The optimal model was chosen to be the random quadratic model that had a common variance-covariance structure across classes because it met all other recommended model adequacy criteria and the posterior classification table of the model indicated that women allocated to the class with an APPA below 70% had a posterior probability of only 17% and 20% of belonging to one of the other classes.

5.4.4 Clinical characterisation and plausibility of identified depressive symptom trajectories

5.4.4.1 Graphical presentation of trajectories

The favoured model identified three groups with different patterns of depressive symptom over time (Figure 5.4). Most women were allocated to a class with stable low depressive symptoms (n = 8,118, 68.8%), followed by the class with stable moderate depressive symptoms (n = 3,558, 30.1%), and the class with fluctuating depressive symptoms (n = 128, 1.1%).

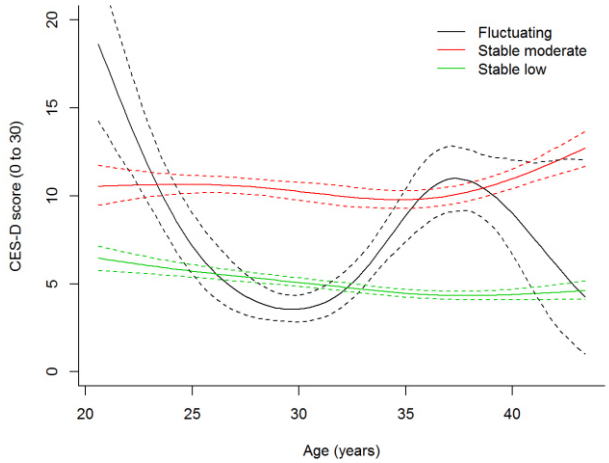


Figure 5.4: Class-specific mean predicted depressive symptoms trajectories and 95% prediction intervals

Whilst depressive symptom scores of women allocated to the stable low class were generally lower than the scores of women allocated to the stable moderate class, as expected there was some overlap in depressive symptom scores of these classes (Figure 5.5). Furthermore, there was variation between individuals allocated to the same class. The individual depressive symptom trajectories of those allocated to the fluctuating depressive symptoms class closely approximated the mean predicted trajectories of that class.

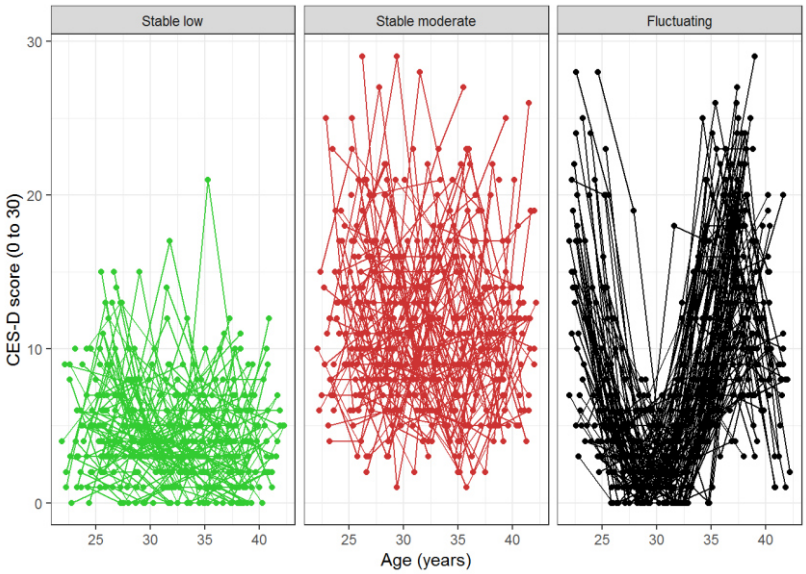


Figure 5.5: Depressive symptom trajectories of 100 randomly selected women of each of the three identified latent classes

5.4.4.2 Profiles of key cardiovascular risk factors among women with different depressive symptom trajectories

5.4.4.2.1 Sociodemographic characteristics and health behaviours

Using data from the subgroup of women that participated at baseline and end of follow-up (n = 7,149, 60.6%), there were statistically significant differences between groups in all sociodemographic characteristics and health behaviours except for area of residence (Table 5.7, Table 5.8). At baseline, women with stable low depressive symptoms were less likely to be separated, divorced or widowed, to live alone, and found it easier to manage on their income. Furthermore, the proportion of current smokers, risky/ high risk alcohol intake, and nil/ sedentary or low physical activity levels was lowest among this group. Women with stable moderate depressive symptoms reported lower educational attainment and found it harder to manage on their income than women allocated to the other two groups. Whilst women with stable moderate depressive symptoms reported higher smoking prevalence than women with fluctuating depressive symptoms, differences between these two groups were small with regard to risky/ high risk alcohol intake. The proportion of women with nil/ sedentary or low physical activity levels was highest among those with fluctuating depressive symptoms.

*Table 5.7: Baseline characteristics, separately for women with stable low, stable moderate and fluctuating depressive symptoms over time**

	Stable low (n = 4,285)	**Stable moderate** (n= 1,701)	**Fluctuating** (n = 100)	**p**
Age (median [IQR])	24.6 [23.3, 25.8]	24.6 [23.4, 25.8]	24.1 [22.8, 25.3]	0.023
Marital status (%)				0.001
Married/ de-facto	1,880 (44.2)	747 (43.9)	46 (46.0)	
Separated/ divorced/ widowed	36 (0.8)	36 (2.1)	1 (1.0)	
Single	2,327 (54.7)	911 (53.6)	51 (51.0)	
Lives alone (%)	269 (6.3)	135 (7.9)	7 (7.0)	0.277
Area of residence (%)				0.859
Metropolitan	2,414 (56.7)	978 (57.5)	63 (63.0)	
Rural	1,675 (39.3)	659 (38.7)	35 (35.0)	
Remote	150 (3.5)	56 (3.3)	2 (2.0)	
University degree (%)	2,131 (50.0)	627 (36.9)	44 (44.0)	<0.001
Ability to manage on income (%)				<0.001
Difficult/ impossible	540 (12.7)	388 (22.8)	14 (14.0)	
Difficult some of the time	1,266 (29.7)	586 (34.5)	33 (33.0)	
Easy/ not too bad	2,438 (57.3)	721 (42.4)	53 (53.0)	
Alcohol intake (%)				<0.001
Low risk	2,736 (64.3)	933 (54.9)	59 (59.0)	
Non/ rarely	1,377 (32.3)	661 (38.9)	35 (35.0)	
Risky/ high risk	126 (3.0)	92 (5.4)	6 (6.0)	
Smoking status (%)				<0.001
Never-smoker	2,707 (63.6)	900 (52.9)	53 (53.0)	
Ex-smoker	560 (13.2)	264 (15.5)	20 (20.0)	
Current smoker	964 (22.6)	521 (30.6)	26 (26.0)	
Physical activity (%)				<0.001
Nil/ sedentary	319 (7.5)	190 (11.2)	10 (10.0)	
Low	1,280 (30.1)	566 (33.3)	41 (41.0)	
Moderate	1,043 (24.5)	412 (24.2)	16 (16.0)	
High	1,603 (37.6)	530 (31.2)	33 (33.0)	

* Data are given as n (%) unless specified; IQR: interquartile range

In keeping with differences observed at baseline, the proportion of women who had a university degree and reported that it was easy/ not too bad to manage on their income was highest and the proportion of women who were separated, divorced, or widowed was lowest among those with stable low depressive symptoms (Table 5.8). Furthermore, the proportion of women with a university degree and who found it easy/ not too bad to manage on their income was lowest among those with stable moderate depressive symptoms. Adverse health behaviours were least prevalent in

women with stable low depressive symptoms at the end of follow-up. The proportion of women who were current smokers and physically less active was highest among those with stable moderate depressive symptoms. Risky/ high risk alcohol intake was most prevalent among those with fluctuating depressive symptoms.

*Table 5.8: Participant characteristics at study phase 7, separately for women with stable low, stable moderate, and fluctuating depressive symptoms over time**

	Stable low (n = 4,285)	**Stable moderate** (n = 1,701)	**Fluctuating** (n = 100)	**p**
Age (median [IQR])	39.6 [38.4, 40.9]	39.7 [38.5, 40.9]	39.2 [38.0, 40.3]	0.012
Marital status (%)				<0.001
Married/ de-facto	3,432 (80.6)	1,101 (64.7)	70 (70.0)	
Separated/ divorced/ widowed	263 (6.2)	195 (11.5)	12 (12.0)	
Single	381 (8.9)	301 (17.7)	11 (11.0)	
Lives alone (%)	223 (5.2)	161 (9.5)	9 (9.0)	<0.001
Area of residence (%)				0.828
Metropolitan	2,410 (56.6)	947 (55.7)	59 (59.0)	
Rural	1,465 (34.4)	586 (34.5)	32 (32.0)	
Remote	125 (2.9)	49 (2.9)	4 (4.0)	
University degree (%)	2,506 (58.9)	756 (44.4)	55 (55.0)	<0.001
Ability to manage on income (%)				<0.001
Difficult/ impossible	388 (9.1)	421 (24.8)	15 (15.0)	
Difficult some of the time	1,040 (24.4)	530 (31.2)	28 (28.0)	
Easy/ not too bad	2,649 (62.2)	645 (37.9)	50 (50.0)	
Alcohol intake (%)				<0.001
Low risk	2,584 (60.7)	816 (48.0)	44 (44.0)	
Non/ rarely	1,342 (31.5)	682 (40.1)	40 (40.0)	
Risky/ high risk	225 (5.3)	139 (8.2)	14 (14.0)	
Smoking status (%)				<0.001
Never-smoker	2,773 (65.1)	907 (53.3)	52 (52.0)	
Ex-smoker	1,057 (24.8)	484 (28.5)	33 (33.0)	
Current smoker	322 (7.6)	239 (14.1)	13 (13.0)	
Physical activity (%)				<0.001
Nil/ sedentary	542 (12.7)	360 (21.2)	15 (15.0)	
Low	1,132 (26.6)	478 (28.1)	28 (28.0)	
Moderate	911 (21.4)	312 (18.3)	11 (11.0)	
High	1,400 (32.9)	401 (23.6)	37 (37.0)	

* Data are given as n (%) unless specified; IQR: interquartile range

Within each identified depressive symptom subgroups, the characteristics of women differed between baseline and end of follow-up. Among all groups the proportion of women who were separated, divorced or widowed; lived alone; with a university degree; and with risky/ high risk alcohol intake was higher at the end of follow-up than at baseline. In contrast, among all groups the proportion of current smokers was smaller at the end of follow-up. The proportion of women who reported that it was easy/ not too bad to manage on their income increased among those with stable low depressive symptoms whereas it decreased among those with stable moderate or fluctuating depressive symptoms from baseline to the end of follow-up. Whilst the proportion of women who report high physical activity increased among women with fluctuating depressive symptoms, it decreased among women with stable low and stable moderate depressive symptoms.

5.4.4.2.2 Body mass index

There were statistically significant differences between the three groups with regard to body mass index. Both at baseline and end of follow-up the proportion of women with overweight or obesity was lowest among the group with stable low depressive symptoms and highest among the group with stable moderate depressive symptoms (Table 5.9, Table 5.10). Among all groups, the mean BMI was higher and the proportion of women with overweight or obesity was greater at the end of follow-up than at baseline.

*Table 5.9: Body mass index of women with stable low, stable moderate and fluctuating depressive symptoms at baseline**

	Stable low (n = 4,285)	**Stable moderate** (n = 1,701)	**Fluctuating** (n = 100)	**p**
Body mass index (kg/m²) (mean (SD))	23.4 (4.4)	24.8 (5.7)	24.0 (4.8)	<0.001
BMI (WHO groups) (%)				<0.001
Underweight	257 (6.0)	101 (5.9)	6 (6.0)	
Acceptable weight	2,683 (63.0)	884 (52.0)	65 (65.0)	
Overweight	745 (17.5)	357 (21.0)	13 (13.0)	
Obese	340 (8.0)	251 (14.8)	11 (11.0)	

* Data are given as n (%) unless specified; SD: standard deviation

*Table 5.10: Body mass index of women with stable low, stable moderate and fluctuating depressive symptoms at the end of follow-up**

	Stable low (n = 4,285)	**Stable moderate** (n = 1,701)	**Fluctuating** (n = 100)	**p**
Body mass index (kg/m²) (mean (SD))	26.3 (5.9)	28.9 (7.6)	28.1 (7.3)	<0.001
BMI (WHO groups) (%)				<0.001
Underweight	67 (1.6)	33 (1.9)	0 (0.0)	
Acceptable weight	2,025 (47.6)	561 (33.0)	39 (39.0)	
Overweight	1,127 (26.5)	431 (25.3)	31 (31.0)	
Obese	915 (21.5)	605 (35.6)	26 (26.0)	

* Data are given as n (%) unless specified; SD: standard deviation

Among 11,355 women with at least one measure of BMI, there were differences in BMI trajectories among those with stable low, stable moderate, and fluctuating depressive symptoms (Figure 5.6). The mean BMI was consistently lower among women with stable low depressive symptoms but differences between groups were small at the age of 20 years. From the age of 20 to 40 years, the mean BMI increased among all depressive symptom groups. The change of BMI over time was significantly different between classes (Wald: 226.1, $p < 0.0001$). Whilst changes of BMI over time were similar for women with stable moderate and fluctuating depressive symptoms, the increase in mean BMI was less marked among those with stable low depressive symptoms. CIs were wide among those with fluctuating depressive symptoms.

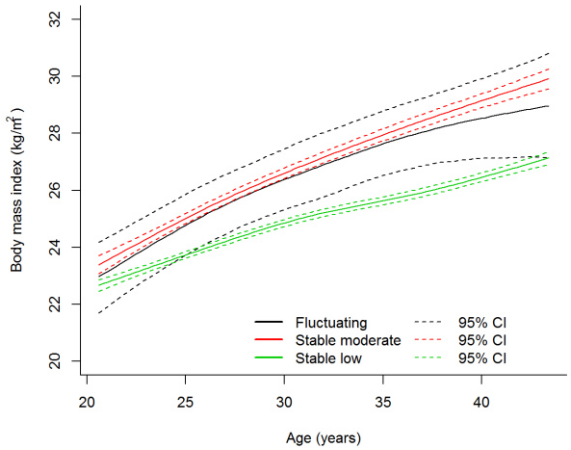


Figure 5.6: Class-specific mean predicted trajectory of body mass index for hypothetical woman with stable low, stable moderate, and fluctuating depressive symptoms over time

5.4.4.2.3 Medical conditions

Among the subgroup of women that participated at baseline and end of follow-up (n = 7,149, 60.6%), the proportions of women with heart disease, diabetes, and hypertension were small among all of the groups (Table 5.11). There were no statistically significant differences between groups with regard to hypertension, diabetes and heart disease at baseline whereas there were statistically significant differences at the end of follow-up (Table 5.11). At the end of follow-up hypertension was most prevalent among those with fluctuating depressive symptoms and lowest among those with stable low depressive symptoms. The proportion of women with diabetes and heart disease was highest among those with stable moderate depressive symptoms. However, this should be interpreted with caution due to the small number of events, especially in the group with fluctuating depressive symptoms.

*Table 5.11: Numbers and proportion of women with hypertension, diabetes, and heart disease, separately for those with stable low, stable moderate, and fluctuating depressive symptoms over time**

	Stable low (n = 4,285)	**Stable moderate** (n = 1,701)	**Fluctuating** (n = 100)	**p**
	Baseline characteristics			
Hypertension (%)	94 (2.2)	59 (3.5)	2 (2.0)	0.064
Diabetes (%)	9 (0.2)	6 (0.4)	0 (0.0)	0.691
Heart disease (%)	2 (0.0)	4 (0.2)	0 (0.0)	0.244
	Characteristics at end of follow-up			
Hypertension (%)	372 (8.7)	240 (14.1)	15 (15.0)	<0.001
Diabetes (%)	108 (2.5)	72 (4.2)	1 (1.0)	0.001
Heart disease (%)	29 (0.7)	27 (1.6)	0 (0.0)	0.003

* Data are given as n (%) unless specified

The results were similar when the proportion of women with hypertension, diabetes, and heart disease was compared among women that were diagnosed with medical conditions at any point during follow-up (Appendix Table A 36).

5.4.5 Sensitivity analyses

5.4.5.1 Restriction of analysis to women with at least two measures of depressive symptoms

When the analysis was restricted to 10,452 women with at least two measures of depressive symptoms, the favoured model was a random slope model with three latent classes. The depressive symptom trajectories of each of these classes were very similar to those identified in the main analysis (Figure 5.7). Of note, the four and five class models resulted in lower discrete AIC values than the three class model; however, these models resulted in one and two classes with less than 1% of all women allocated to the classes, respectively (Appendix Table A 37). Furthermore, the mean predicted trajectories of some classes largely overlapped (Appendix Figure A 2). As a result, three classes were chosen as preferred number of classes.

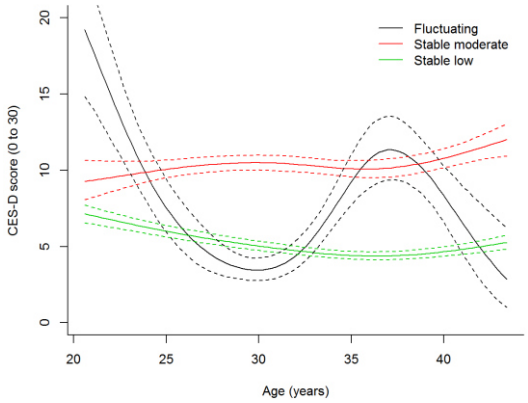


Figure 5.7: Mean predicted depressive symptoms trajectories of three latent classes in analyses restricted to women with at least two assessments of depressive symptoms

5.4.5.2 Unique CES-D items

5.4.5.2.1 Item: I felt depressed

Following the same steps as reported above, the favoured model to describe trajectories of the CES-D item "I felt depressed" was identified. Using the unique item "I felt depressed" as dependent variable, the favoured model was identified as a random slope model with one latent class (Figure 5.8). The predicted mean score of the item "I felt depressed" was low at the age of 20 years, and the mean predicted trajectory declined very slightly over time from 20 to 35 years of age. Although the favoured model was different to that identified in the main analysis, the depressive

symptom trajectories of the one and two class solution looked very similar to those of the main analysis (Appendix Figure A 3). The three and four class solution did not reach convergence in 500 iterations (Appendix Table A 38).

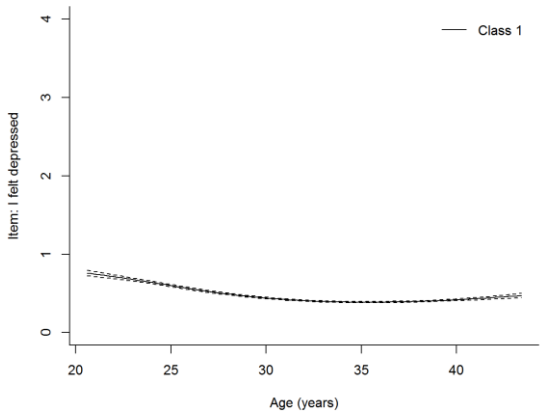


Figure 5.8: Mean predicted trajectory of unique CES-D item "I felt depressed"

5.4.5.2.2 Item: I felt hopeful about the future

The favoured model to describe trajectories of the CES-D item "I felt hopeful" was identified as random quadratic model with one latent class (Figure 5.9). Similar to the trajectory of scores of the item "I felt depressed", the mean predicted score of "I felt hopeful about the future" was low at 20 years of age and declined very slightly over time. Again, the three and four class solution did not reach convergence (Appendix Table A 39), and the one and two class solutions were similar to those observed in the main analysis (Appendix Figure A 4).

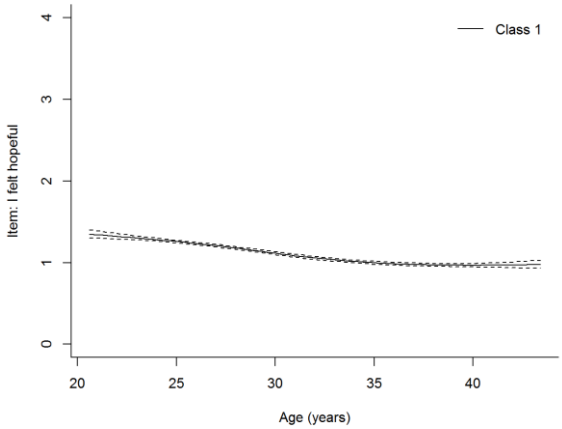


Figure 5.9: Mean predicted trajectory of unique CES-D item "I felt hopeful about the future"

5.5 Discussion

5.5.1 Summary of main findings

Among Australian women three subgroups of women with distinct depressive symptom trajectories were identified using group-based trajectory modelling. From the age of 20 to 40 years the majority of women had stable low depressive symptoms but there was a considerable proportion of women with stable moderate depressive symptoms and a small number of women with fluctuating depressive symptoms.

There were differences between women with different patterns of depressive symptoms as well as differences between baseline and end of follow-up with regard to sociodemographic characteristics, health behaviours and proportions of women with medical conditions. At baseline and end of follow-up women with stable moderate or fluctuating depressive symptoms had lower educational attainment, found it harder to manage on their income, reported more adverse health behaviours, and had more medical conditions than women with stable low depressive symptoms. Among all groups, the proportion of current smokers decreased and proportion of women with risky/ high risk alcohol intake increased from baseline to the end of follow-up. There were differences in the patterns with regard to ability to manage on income and physical activity levels. Although the proportion of women with a university degree increased among all groups, the proportion of women who found it easy/ not too bad to manage on their income decreased among those with stable moderate and fluctuating depressive symptoms whereas it increased among those with stable low depressive symptoms. Furthermore, the proportion of women with high physical activity levels increased among women with fluctuating depressive symptoms but decreased among the other two groups between baseline and end of follow-up. Interestingly, the mean BMI of women with different patterns of depressive symptoms in young adulthood was similar at the age of 20 years. However, the patterns of change of BMI over time differed between women with stable low, stable moderate, and fluctuating depressive symptoms. Although the mean BMI of all groups increased over time, the increase was more gradual among those with stable low depressive symptoms. This might be particularly meaningful

considering that BMI is one of the potential mediators between depression and subsequent CVD (see sections 7.4 Implications for research and 7.5 Implications for practice for more detailed discussion).

5.5.2 Strengths and limitations of this analysis

5.5.2.1 Strengths

This analysis has major strengths. First, it made use of repeat assessments of depressive symptoms using the CES-D rating scale, a well validated rating scale to assess depressive symptoms. The CES-D rating scale was validated against the revised Clinical Interview Schedule (CIS-R), a structured diagnostic interview for common mental disorders (Lewis et al, 1992), using a random sample of Whitehall II participants supplemented with Whitehall II participants with depression (Head et al, 2013). Sensitivity and specificity values for detecting any mental disorder were 77% and 89%. Sensitivity and specificity for detecting depressive episode were 89% and 86% for the CES-D. Head et al (2013) concluded that the CES-D had good validity for detecting any mental disorder and depressive episodes. Furthermore, the shortened 10-item version of the CES-D that was used in this analysis was shown to have good validity and test-retest reliability (Carpenter et al, 1998; Furukawa et al, 1997; Mohebbi et al, 2018). Existing studies often relied on a single measure of depressive symptoms, implicitly assuming that this measure approximates the participant's exposure status throughout follow-up. In contrast to existing studies, this analysis did not assess the risk of CVD among those with and without depressive symptoms at one point in time. Instead, the trajectories of depressive symptoms over 20 years of age were considered and profiles of key cardiovascular risk factors of women with similar depressive symptom trajectories were investigated. The identification of three groups with different patterns of depressive symptoms highlights the importance of using repeated measures of depressive symptoms, especially since there were differences in the profiles of key cardiovascular risk factors across the three trajectory groups at the beginning and end of follow-up.

This is one of the first studies that investigated trajectories of BMI among women with different patterns of depressive symptoms over time. This analysis is of particular interest since BMI has been highlighted as the main potential confounding and/ or mediating factor in the relationship between depression and hypertension in an existing publication using the mid-aged cohort of the ALSWH (Jackson & Mishra, 2013). It will be of interest to further investigate the relationship between depression or depressive symptoms, BMI, and CVD in future studies (see section 7.4 Implications for research for more detailed discussion).

Further strengths of this analysis were that a framework to construct latent class trajectory models was followed, recommended model assessment tools were used to select the favoured model, and that all decisions were described transparently. This was particularly important since differences between existing studies using latent class trajectory modelling might have been due to different modelling assumptions rather than true differences between populations (Lennon et al, 2018). However, it should be noted that, although this analysis followed recommended steps of the framework to construct latent class trajectory models, the selection of the favoured model involved making a number of decisions that were largely subjective. For example, although the models with five and four number of groups had lower discrete AIC values than the model with three groups, the model with three groups was selected as the preferred model since it identified larger subgroups and showed non-overlapping depressive symptom trajectories. Other researchers might disagree with this decision. As a result, the favoured model of other researchers might have differed from the one identified in this analysis. However, all decisions were described transparently and an attempt was made to balance criteria of model fit and clinical usefulness.

In contrast to existing studies, the model selection criterion that was chosen in this project allowed for specific properties of the depressive symptoms measure. As described previously (see section 5.3.3.2.2 Link function), the discrete AIC takes into account that the measure of depressive symptoms is a discrete and bounded

quantitative outcome by using the discrete log-likelihood instead of the continuous log-likelihood. To the best of my knowledge, no discrete BIC has been described in the literature. In the context of latent class mixed modelling most studies have based their model selection on the BIC. Despite recommendations to use the BIC as model selection tool (Lennon et al, 2018; van de Schoot et al, 2017), I decided that taking into account specific properties of the depressive symptom measure outweighed the importance of a more general recommendation of an expert committee. A further model selection tool to choose the preferred number of classes is the Lo-Mendel-Rubin-likelihood ratio test (Lo et al, 2001). This model selection tool is based on a hypothesis test. If $p < 0.05$, the model with fewer classes will be rejected in favour of a model with at least one more class. I decided not to use this model selection tool since van de Schoot et al (2017) reported that simulation studies suggested that the results may not be correct. Instead, I determined the optimal number of classes balancing the result of the chosen statistical tool (lowest discrete AIC) against criteria reflecting the clinical relevance and usefulness of the identified subgroups (e.g. group size). However, I acknowledge that there are several statistical criteria that could have been used to choose the preferred model and selecting the discrete AIC might have influenced the selection of the preferred model.

5.5.2.2 Limitations

5.5.2.2.1 Chance

This analysis might be influenced by potential sources of error. A well-recognised disadvantage of group based trajectory modelling is the potential lack of statistical power due to differently sized subgroups and small numbers of individuals in each of the latent classes (Vistisen et al, 2014). Whilst the groups with stable low and stable moderate depressive symptoms were relatively large, the number of women with fluctuating depressive symptoms was small. As a result, there was limited power to detect statistically significant differences between this class and the other classes. Furthermore, the CIs of the mean predicted depressive symptom trajectory among those with fluctuating depressive symptoms were wide after the age of 37 years, indicating imprecise estimation. There was an apparent peak of depressive symptoms

in the mid to late thirties and a downward trend thereafter among women with fluctuating depressive symptoms. To the best of my knowledge, no other study has reported this pattern. This might be due to differences in model specifications, differences across samples, or differences across mental health measures. However, another explanation is that the apparent peak in the mid to late thirties and apparent downward trend in this sample was observed due to chance alone.

The data requirements with regards to sample size and number of data points (see section 5.3.3.2.1 Latent process mixed modelling) were met in this analysis. Furthermore, the sample size of this analysis was bigger than 22 out of 25 existing studies that were identified in a systematic review of studies using group-based trajectory models (Musliner et al, 2016). Of the three studies with a similar or bigger sample size, one study was based on three data points (Costello et al, 2008), one study was based on five data points (Melchior et al, 2013), and one study was based on up to six data points (Liang et al, 2011). A comparison of the statistical power of this analysis with existing studies is complicated since statistical power is not only affected by the sample size and number of data points but also by the complexity of the model structure. Generally, there is a trade-off between statistical power and the complexity of the model structure. For example, the decision to select the model with three groups as preferred model had an impact on subsequent steps of the analysis. When five models with increasingly complex model structures were fitted, the model with quadratic random effects and a variance-covariance matrix that varied across classes with a proportionality constraint did not converge. This was likely driven by limited statistical power due to the small trajectory group of women with fluctuating depressive symptoms in combination with the complexity of the model structure. To investigate this further, I performed a post-hoc analysis in which I selected the preferred number of classes as two. I fitted models with increasingly complex model structures as described in Table 5.1. In this post-hoc analysis, all models reached convergence in 500 iterations. This highlights the need to carefully balance statistical criteria and criteria such as clinical plausibility and meaningfulness of trajectory groups, especially in the context of small trajectory groups.

Due to the young age of women at the end of follow-up (37 – 42 years), the number of medical conditions was small. As explained previously (see section 5.1 Background), the young cohort of the ALSWH study was chosen because at the outset of this project one aim of the analysis was to investigate the potential mediating role of BMI in the relationship between depression and hypertension. Since an existing study based on the mid-aged cohort of the ALSWH concluded that the vast majority of women were already overweight at baseline, it was of interest to investigate the relationship further using the young cohort of the ALSWH. It will be of interest to further investigate the relationship between different patterns of change of depressive symptoms and risk of CVD in the young cohort of the ASLWH once women are older and more events have occurred (see section 7.4 Implications for research for more detailed discussion).

5.5.2.2.2 Bias

Women might have been classified to a depressive symptom trajectory group that did not accurately represent their course of depressive symptoms from the age of 20 to 40 years. First, assessments of depressive symptoms were based on self-report which might be influenced by under- or over-reporting (Schubert et al, 2017). However, considering the large number of women enrolled in the ALSWH study and the number of follow-up assessments, interviews to ascertain clinical depression would not have been feasible. Also, potentially important differences in levels of subthreshold depression would remain undetected if a measure of clinical depression was used. Second, class membership was determined using the highest posterior group membership probability. Allocating women to the class with the highest posterior class membership probability assumes that class allocation was possible without classification error. However, it has been highlighted by other researchers that there might be uncertainty about class allocation (van de Schoot et al, 2017). Entropy values can guide an assessment of the extent of bias due to classification error in that higher entropy values indicate fewer classification errors and less bias in the prediction of class membership. Although the relative entropy value exceeded the

threshold of 0.5 recommended by Lennon et al (2018), it only did so by a small margin in this analysis.

Due to the use of a forward timescale, events during the course of the trajectory might have altered the trajectory itself. For example, if a women was diagnosed with hypertension at the age of 30 years, this might have altered her depressive symptom trajectory from the age of 30 to 40 years. The use of a forward timescale and age as a marker of time is useful if researchers are interested in the natural course of a marker over time. In contrast, the use of a backward timescale is warranted if researchers are interested in depressive symptom trajectories after a significant life event or before the diagnosis of medical conditions (Vistisen & Færch, 2014). Furthermore, a backward timescale is useful if one is interested in the aetiology of disease development (Vistisen & Færch, 2014) or for scenarios in which reverse causation might explain the observed relationship between the exposure and outcome of interest. Since the observed association between depression or depressive symptoms and subsequent CVD might be explained by reverse causation, group-based trajectory modelling using a backward timescale will be further explored in the next chapter (see Chapter 7).

In contrast to existing studies, women with only one measurement of depressive symptoms were not excluded from the main analysis. Since women with only one assessment of depressive symptoms might be systematically different from women with more than one assessment of depressive symptoms, an exclusion of these women might introduce selection bias (Boucquemont et al, 2017). When investigating individual depressive symptom trajectories of a random subset of women included in the analysis, there was some indication that women with only one assessment of depressive symptoms had higher depressive symptom scores (see section 5.3.3.1 Descriptive statistics). Exclusion of these women might therefore have led to an underestimation of the mean score of depressive symptoms in this population. Since existing studies have frequently excluded women with only one measurement of depressive symptoms (Leigh et al, 2016; Nabi et al, 2008; Tran et al, 2019), a sensitivity

analysis was performed among women with at least two assessment of depressive symptoms during follow-up. In this analysis, the best model identified three latent classes that were very similar to those identified in the main analysis. Despite including women with only one assessment of depressive symptoms, there might have been some selection bias due to the exclusion of women who only participated in study phase 1 or who did not have any depressive symptom assessment. Unfortunately, differences between those included and excluded could not be assessed with regard to depressive symptoms since the CES-D scale was first introduced in study phase 2. Subjects with missing data on depressive symptoms were included in the analysis, but only available data for each subject were used in the analysis. This generated unbiased parameter estimates under the assumption that data were MAR (Nagin & Odgers, 2010). MAR means that missingness is conditional on observed variables but unconditional on unobserved variables. Variables that were observed and used in the trajectory analysis were depressive symptoms, age, sex (through restriction to women), and study wave. Neither age, sex, nor study wave had any missing data. At each wave of data collection the proportion of women with missing information on depressive symptoms was relatively low (see Table 5.2). Furthermore, missingness of depressive symptoms is assumed to be random given the observed measurements of depressive symptoms. Therefore, it seems likely that the MAR is valid. However, there might have been some bias due to the exclusion of women without any measurement of depressive symptoms throughout follow-up or due to women with few measurements of depressive symptoms.

The sample included Australian women who were randomly selected from the national Medicare health insurance database. The response rate to the initial invitation letter was estimated at 41 – 42% for the young cohort with some uncertainty remaining due to unknown accuracy of the Medicare records (Brown et al, 1999). Due to data confidentiality, demographic characteristics of responders and non-responders could not be compared. Based on aggregate data, there was some indication that non-responders made less use of health services (Brown et al, 1999). Comparing participant characteristics at baseline to 1996 Australian census data, it

has been shown that the sample were reasonably representative of general Australian population, although well-educated women were overrepresented and immigrant women underrepresented (Brown et al, 1999). The extent to which the trajectories of depressive symptoms and their association with cardiovascular risk are generalisable to women from countries other than Australia, to older women, and to men has yet to be established in other datasets with multiple measures of depressive symptoms over time. As described previously, women from rural areas were intentionally oversampled. This would affect results of the analysis if the prevalence and pattern of trajectories of depressive symptoms over time was different among women from rural and urban areas. Whilst prior research showed evidence for rural-urban differences with regard to schizophrenia and suicides, evidence with regard to common mental health disorders such as depression and anxiety was inconclusive (Solmi et al, 2017). Thus, it remains to be determined to what extent the oversampling of women from rural areas affected the prevalence and/ or pattern of each of the depressive symptom trajectories.

The use of a depressive symptom sum score might have biased the analysis due to a violation of the measurement invariance and unidimensionality assumptions. As discussed in section 5.3.3.4 Sensitivity analysis, previous research has shown that depressive symptom sum scores might not measure one underlying construct (unidimensionality) and might not measure that construct the same way across time (measurement invariance) (Fried et al, 2016; van Eeden et al, 2019). A violation of these assumptions changes the interpretations of the findings. Changes in the sum score would not reflect depressive symptom severity but rather have to be interpreted as a change in the sum of different underlying problems (Fried et al, 2016). Since the pattern of change of a sum of different underlying problems over time might not be meaningful, researchers have recommended to put more emphasis on the study of individual items (Fried et al, 2016; van Eeden et al, 2019). I followed this recommendation by modelling the pattern of change of two unique items of the depressive symptom scale over time. These unique items followed similar patterns over time as the depressive symptom sum score. However, for both unique items the

preferred models identified a smaller number of classes. This might be explained a lack of statistical power since the unique items ranged from 0 to 4 whereas the sum score ranged from 0 to 30. The results of the sensitivity analysis should be interpreted with caution because the CES-D scale was not developed to analyse individual items. Fried et al (2016) recommended to increase the reliability of symptom measurement by using multiple items to measure the same unique symptoms. However, this was not possible using the CES-D scale because the scale assessed each unique symptom using one item.

5.5.2.2.3 Confounding

Unconditional latent process mixed models were fitted that explored patterns of depressive symptoms among Australian women from the age of 20 and 40 years. As such, the number of latent classes in this sample were explored without consideration of covariates. After allocation of women to the class with the highest posterior class membership probabilities, potential predictors of class membership and characteristics at the end of follow-up were explored. van de Schoot et al (2017) highlighted that researchers should not combine the step of finding the number of relevant classes with the step of selecting the main predictors of those classes. One of the reasons for this is that the inclusion of predictors overestimates the entropy value and thereby underestimates classification uncertainty.

The comparison of participant characteristics of the three depressive symptom subgroups at baseline and end of follow-up should be interpreted with caution since they were not adjusted for potential confounding factors. I intended to perform logistic regression model to assess the association between class membership and sociodemographic characteristics, health behaviours, and variables related to the health status of participants. However, after identifying the favoured latent process mixed model, I decided to describe and discuss differences between groups rather than to run a regression model since the small number of women with fluctuating depressive symptoms would have resulted in imprecise estimates.

5.5.3 Comparison of findings with previous research

Using data from the youngest and oldest cohort of the ALSWH, existing studies have investigated trajectories of mental health but not depressive symptoms. Using logistic group-based trajectory modelling, Holden et al (2016) identified four groups with different patterns of mental health in the young cohort of the ALSWH with complete data available. Most participants were allocated to a trajectory group that was characterised by stable high mental health (55%), whilst 24%, 12%, and 9% were allocated to groups with varying, improving, and stable low mental health over 16 years of follow-up, respectively. In contrast to the analysis presented in this chapter, Holden et al (2016) did not use a measure of depressive symptoms but the 5-item Mental Health index which is a non-specific measure of mental health-related QOL. Furthermore, the authors dichotomised their measure to identify individuals with low versus high mental health throughout follow-up because of which potentially important information on variations in scores of mental health over time were missed. Another disadvantage of the study by Holden et al (2016) is that it was based on complete cases which might have introduced selection bias. In a recently published paper based on the young cohort of the ALSWH, Xu et al (2019) investigated if trajectories of mental health among women in their 20s predicted chronic physical conditions in their 30s. Chronic physical conditions were defined as diabetes, cancer, hypertension, CVD, asthma, and bronchitis. Using the 5-item Mental Health Index, the authors identified five mental health trajectories of high-stable (60%), improving (9.5%), declining then improving (19.2%), declining (4.8%), and low stable (6.4%) mental health-related QOL. Relative to women with high stable mental health, women allocated to the low stable, declining, declining then improving, and improving trajectory groups had higher odds of chronic physical condition. CIs overlapped with the null value in the latter two groups. Whilst the association between trajectories of mental health of women in their 20s with incidence of chronic physical diseases of women in their 30s is interesting, a disadvantage of the analysis is that about 3,000 women with chronic physical conditions in their 20s were excluded. This approach might have introduced selection bias and arbitrarily

separates the follow-up period into an exposure window (in which the trajectories are identified) and an outcome window (in which the odds of developing chronic physical conditions are estimated). Using data from the oldest cohort of the ALSWH (aged 70 – 75 years at baseline), two studies identified three different patterns of trajectories of mental health among women as they age (Leigh et al, 2016; Tran et al, 2019). Whilst the number of groups was the same in both studies, the pattern of trajectories and proportion of women allocated to each trajectories differed between the two studies. Leigh et al (2016) used the 5-item Mental Health Index to allocate 28%, 44% and 28% to groups characterised by excellent, good, and poor mental health, respectively. The trajectories were relatively flat with little change over time, although there was a slight decrease in mean mental health scores among those with poor mental health. Tran et al (2019) used the 4-item SF36 Vitality subscale which is a measure of positive mental health and identified a larger proportion of women with stable high mental health score (77%), a small group with declining mental health (4.8%), and 18.2% with stable low mental health. Differences between studies might be explained by the use of different mental health measures.

Musliner et al (2016) published a systematic review on patterns, predictors and outcomes of long-term trajectories of depressive symptoms in studies using group-based trajectory modelling. The authors identified 25 studies, seven of which were based on children or adolescents, four of which were of young adults (< 60 years), eight of older adults, and six studies were based on mothers. The identified number of latent classes ranged from three to six, with most studies identifying either three or four latent classes. The classes varied in their symptom severity and symptom stability over time. In most studies the largest class was characterised by stable low depressive symptoms. Other classes were characterised by persistently high depressive symptoms, stable moderate depressive symptoms, and unstable depressive symptoms, either increasing or decreasing. These findings were in keeping with another systematic review on depressive symptom trajectories in late adolescence and early adulthood (Schubert et al, 2017). In keeping with the majority of existing studies, this analysis identified three subgroups with different patterns of

depressive symptoms over time of which the group with stable low depressive symptoms was the largest. Musliner et al (2016) highlighted that patterns of trajectories were similar for men and women but the proportion of individuals with stable high depressive symptoms was larger among women than men. Furthermore, female gender was a strong predictor of belonging to the class that was characterised by stable high depressive symptoms. Other predictors of stable high depressive symptoms were low income, low education, non-white race, and stressful life events. This was in keeping with the current analysis showing differences between groups in terms of marital status, university degree and difficulty managing on their income.

Despite the potentially important role of BMI as confounding and/ or mediating factor, only few existing studies have investigated BMI trajectories across different mental health subgroups. Gaysina et al (2011) investigated BMI trajectories among men and women with different patterns of affective disorders from adolescence to adulthood (15 to 53 years). Women with adolescent onset affective symptoms had a lower BMI at age 15, a faster increase of BMI over time, and a higher BMI at age 53 years than those without symptoms. Men with adolescent onset affective symptoms had a lower BMI at all ages than men without such symptoms. There was no difference between trajectories of BMI among those with adult onset affective symptoms versus no symptoms. Mumford et al (2013) identified five concurrent trajectories of mental health (MHI-5) and BMI among individuals aged 15 to 27 years of age. Individuals were characterised as having stable good mental health and stable normal weight during follow-up (82.2%), stable good mental health and consistently obese (6.8%), declining mental health and overweight becoming obese (5.6%), improving mental health and stable normal weight (3.3%), and stable good mental health and morbid obesity (2.1%). In keeping with results of this analysis, existing studies reported that the average weight increases over time with differences in patterns of weight gain across groups. Gaysina et al (2011) and Mumford et al (2013) additionally observed differences in the starting point of weight gain whereas the starting point of weight gain was similar among groups in this analysis. Likely

explanations for this difference across studies are differences in the mean age and age range of the samples at baseline.

5.6 Conclusion

This analysis illustrated that group-based trajectory modelling is a useful tool to study patterns of changes of depressive symptoms over time. I identified potentially clinically meaningful subgroups of women with different patterns of depressive symptoms, characterised by differences in profiles of key cardiovascular risk factors across subgroups and differences between baseline and end of follow-up.

Among Australian women three subgroups of women with distinct depressive symptom trajectories were identified from the age of 20 to 40 years. Whilst the majority of women had stable low depressive symptoms, there was a considerable proportion of women with stable moderate depressive symptoms and a small number of women with fluctuating depressive symptoms. Women with stable low, stable moderate, and fluctuating depressive symptoms differed in their profiles of key cardiovascular risk factors. Women with stable moderate or fluctuating depressive symptoms reported lower educational attainment, found it harder to manage on their income, reported more adverse health behaviours, had more medical conditions and gained weight more rapidly than women with stable low depressive symptoms.

Using a forward timescale, this analysis identified women with similar patterns of depressive symptoms in early adulthood. As mentioned previously (see section 5.5.2.2.2 Bias), a retrospective timescale, such as time before the diagnosis of CVD, is useful if one is interested in the aetiology of disease development (Vistisen & Færch, 2014). It can also be valuable in scenarios in which reverse causation might explain the observed relationship between the exposure and outcome of interest. Thus, group-based trajectory modelling using a retrospective timescale will be further explored in the next chapter.

Chapter 6: Trajectories of psychological distress prior to diagnosis of coronary heart disease and stroke

6.1 Background

There are potential alternative explanations for the observed association between clinical depression, depressive symptoms and risk of cardiovascular events. First, as discussed in more detail previously (see section 2.3.2.3 Proposed mechanisms), underlying vascular disease may influence the development of depressive symptoms (vascular depression hypothesis) (Alexopoulos et al, 1997; Taylor et al, 2013). If depressive symptoms are indeed related to worsening vascular disease, depression might not be causally related with subsequent cardiovascular events but instead reflect a prodromal feature of worsening vascular disease itself. Second, researchers have highlighted that a potential explanation for the association between clinical depression, depressive symptoms and risk of cardiovascular events is that symptoms of subclinical CVD overlap with somatic symptoms of depression. If subclinical CVD increase the risk of being diagnosed with clinical depression or depressive symptoms, the apparent increased risk of cardiovascular events among those with clinical depression or depressive symptoms could be due reverse causation. If the association between clinical depression, depressive symptoms and risk of cardiovascular events was indeed due to worsening of underlying vascular disease or the overlap of symptoms, one would expect an increase in depressive symptoms prior to diagnosis of cardiovascular events. As a result, there would be differences between the trajectories of depressive symptoms among those with and without cardiovascular events in the years immediately prior to diagnosis with CVD. No difference between trajectories would be expected in the years prior to expected onset of subclinical CVD.

The Whitehall II study offered unique advantages in investigating trajectories of depressive symptoms prior to diagnosis of cardiovascular events. First, participants of the Whitehall II study have been followed for 28 years. Due to the long subclinical phase of CVD, a long follow-up was needed to cover a time before likely onset of subclinical CVD. Second, the Whitehall II study is one of only a few cohort studies

with repeat assessments of psychological distress and depressive symptoms using the same assessment tools throughout follow-up. A disadvantage of the Whitehall II study is that depressive symptoms were first assessed in study phase 7 whereas psychological distress was assessed in all study phases except for phase 4. It was decided that the benefits of using the more extensive repeated measures of psychological distress outweighed the disadvantages of using an exposure assessment that is less comparable to those used in other parts of this project. However, it should be noted that although the terms depressive symptoms and psychological distress are often used interchangeably in existing literature, psychological distress is a less specific measure of common mental health problems. Whilst some psychological distress scales have a stronger focus on symptoms of depression and anxiety (e.g. Kessler Psychological Distress scale), others capture a wider spectrum of mental health symptoms such as symptoms of social withdrawal (e.g. General Health Questionnaire (GHQ)) (Jackson, 2007).

Existing research has not only shown an association between clinical depression, depressive symptoms and risk of cardiovascular events but also between psychological distress and risk of cardiovascular events. Four studies on the risk of stroke or myocardial infarction among individuals with psychological distress relative to those without psychological distress were identified in the systematic review and meta-analysis performed as part of this project (Brunner et al, 2014; Jackson et al, 2018; Pan et al, 2011a; Whang et al, 2009). One of the studies was performed in the Whitehall II cohort (Brunner et al, 2014). Whilst the search strategy of the systematic review did not include terms for psychological distress, studies using a measure of psychological distress were picked up in the search if they mentioned words starting with depressi* or dysthymi* anywhere in the article. Since it is very likely that studies on psychological distress mention terms such as depression, depressive, or dysthymia at least once, it is very likely that most eligible studies on psychological distress were identified in the literature search. Three of the identified studies assessed psychological distress using rating scales (Brunner et al, 2014; Jackson et al, 2018; Whang et al, 2009) and one study defined the exposure as

psychological distress using a rating scale, physician diagnosis of clinical depression or antidepressant use (Pan et al, 2011a). Using the same methodology as described in Chapter 3, I performed a random effects meta-analysis of these studies. Psychological distress was associated with increased hazards of total stroke, ischaemic stroke, haemorrhagic stroke, and MI (Figure 6.1). The pooled HRs ranged from a 22% increased hazard of MI to a 36% increased hazard of total stroke and haemorrhagic stroke among those with psychological distress, relative to those without psychological distress. There was some evidence of potential age and sex differences in the study by Jackson et al (2018). Furthermore, Brunner et al (2014) highlighted that an association between psychological distress and stroke was observed only over 0 – 5 years but not over 5 – 10 years of follow-up of Whitehall II study participants suggesting that the relationship between psychological distress and stroke might be due to reverse causation.

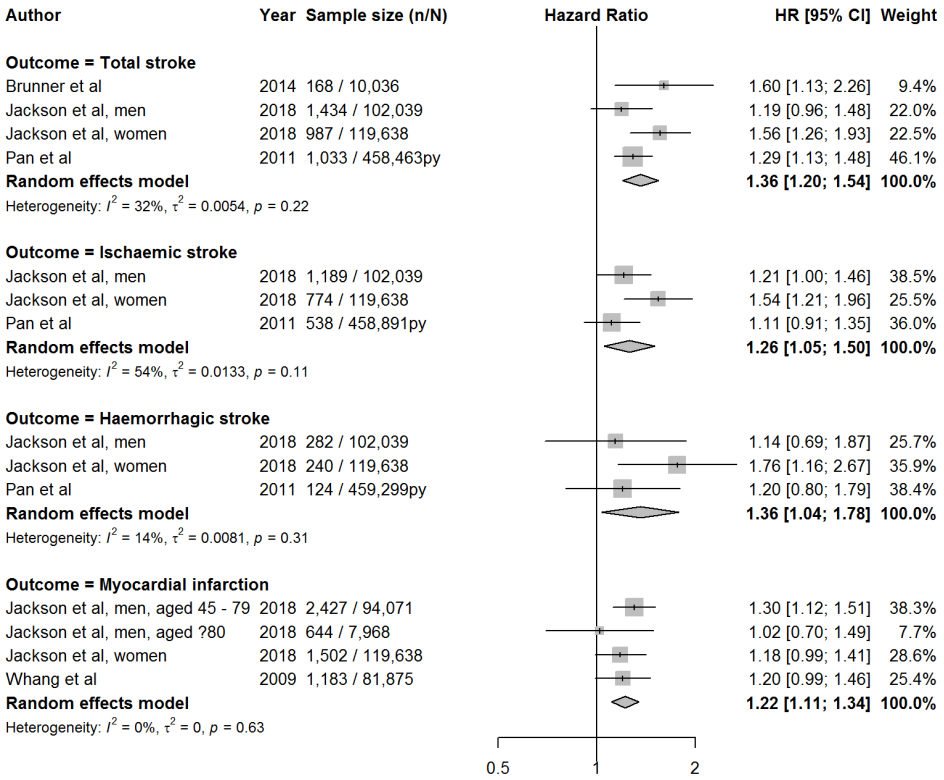


Figure 6.1: Meta-analysis of the risk of total stroke, ischaemic stroke, haemorrhagic stroke and myocardial infarction among individuals with high psychological distress, relative to non-exposed individuals

As highlighted above, existing studies have repeatedly highlighted that the apparent increased risk of cardiovascular events among individuals with clinical depression, depressive symptoms, and psychological distress might be due to reverse causation or overlap in somatic symptoms. Therefore, the aim of the work presented in this chapter was to compare psychological distress trajectories of individuals prior to diagnosis of cardiovascular events to psychological distress trajectories of individuals free from cardiovascular events over the same time period using data from the Whitehall II study. A secondary objective was to investigate to what extent the group of people with cardiovascular events was composed of multiple latent classes that differed in their psychological distress trajectories before diagnosis of cardiovascular events.

6.2 Objectives

1. To compare trajectories of psychological distress of individuals prior to diagnosis of cardiovascular events to individuals free from cardiovascular events over the same period using group-based trajectory modelling
2. To identify homogenous psychological distress subgroups within the heterogeneous group of individuals with a cardiovascular event

6.3 Methods

6.3.1 Study population

The Whitehall II study is an ongoing prospective cohort study of non-industrial British civil servants in 20 London-based departments, aged 35 – 55 years at baseline (Marmot & Brunner, 2005). Individuals were invited by letter and 10,308 individuals (6,895 men and 3,413 women) of mainly white ethnicity were recruited between 1985 and 1988 (73% of those invited). Participants were followed at ten subsequent study phases approximately 2.5 years apart. In all study phases individuals were asked to respond to a questionnaire, and every second phase included a clinical health examination. Participant consent was renewed at each contact and research ethics approval was given by the University College London ethics committee.

6.3.2 Sample

Whitehall II participants were excluded from this analysis if they did not have any assessment of psychological distress or if their record could not be linked with mortality records. Participants with missing data on psychological distress in some but not all study phases were included but only available data for each participant were used in the trajectory analysis. Furthermore, observations for participants were excluded after the occurrence of first-ever fatal CHD or non-fatal MI and fatal or non-fatal stroke or if they contained missing information on person years at risk. Information of study phases with no validated information on outcomes were excluded from the analysis, since the Whitehall II study team advised me not to combine self-reported information on fatal CHD or non-fatal MI and fatal or non-fatal stroke with validated information on events of earlier study phases (see section 6.3.3.2 Fatal CHD/ non-fatal MI and fatal or non-fatal stroke for more detail).

6.3.3 Variables

6.3.3.1 Psychological distress

Psychological distress was self-reported at study phases 1 to 3 and 5 to 11 using the 30-item General Health Questionnaire (GHQ), a screening questionnaire that was developed to identify individuals at high risk of non-psychotic psychiatric illness in the general population (Goldberg, 1972). Participants were asked to rate how their health has been in general over the past few weeks. There were four response options for each of the 30 questions. The overall GHQ score was calculated by the Whitehall II study team using a binary scoring method in which a score of 0 was allocated to the two least symptomatic response options and a score of 1 was allocated to the two most symptomatic response options (Jackson, 2007). As a result, the total score ranged from 0 to 30. The GHQ score was treated as continuous variable in the analyses. In addition, a binary variable indicating high versus low psychological distress was created. In keeping with recommendation, a cut-off of at least five was used to identify individuals with high psychological distress (Jackson, 2007). The binary variable was only used for descriptive statistics.

6.3.3.2 Fatal CHD/ non-fatal MI and fatal or non-fatal stroke

Outcomes were defined as first-ever fatal CHD or non-fatal MI and first-ever fatal or non-fatal stroke. The cause and date of death were ascertained using linkage to the national mortality register kept by the NHS Central Registry. In accordance with Whitehall II procedures, CHD mortality was defined by ICD-9 codes 410 – 414 or ICD-10 codes I20 – I25 and stroke mortality by ICD-9 codes 430 – 438 or ICD-10 codes I60 – 69 (Brunner et al, 2014). Non-fatal CHD and stroke events were validated using NHS Hospital Episode Statistics discharge diagnoses or General Practitioner's (GP) confirmation. Two trained coders independently classified non-fatal events, with adjudication in case of disagreement. Validation of fatal CHD or non-fatal MI and fatal or non-fatal stroke events was available up to study phases 9 and 8, respectively.

6.3.3.3 Covariates

Sociodemographic factors included age, sex, ethnicity, marital status, and occupational position. Ethnicity was grouped into two categories of white and non-white. In keeping with a previous Whitehall II publication, marital status was grouped into two categories of married/ cohabiting and other (Singh-Manoux et al, 2017). Occupational position was ascertained as variable with three categories (administrative, professional/ executive, and clerical/ support workers) based on the participant's salary, social status, and level of responsibility at work (Marmot et al, 1991). Health behaviours included smoking status, alcohol consumption, physical activity, and frequency of fruit and vegetable consumption. Smoking status was grouped into three categories of current smoker, ex-smoker, and non-smoker. In keeping with weekly intake guidelines of the UK Department of Health, alcohol intake was defined as safe drinking if men or women reported ≤ 14 units of alcohol per week and as risky drinking if alcohol intake exceeded 14 units per week (UK Chief Medical Officer, 2016). Physical activity was defined as average hours per week of moderate and vigorous physical activity. In keeping with a previous Whitehall II publication, fruit and vegetable intake was grouped into three categories of less than once a day, daily, and more than once a day (Singh-Manoux et al, 2017). Variables related to the health status of participants included BMI, diabetes, blood pressure,

and total cholesterol levels. BMI (kg/ m^2) was calculated by the Whitehall II study team based on measured height and weight. A history of diabetes at baseline was defined based on self-reported information of a physician diagnosis or diabetes medication. Systolic and diastolic blood pressure were measured in a sitting position using Hawksley random-zero sphygmomanometer after five minutes of rest. Two measures of systolic and diastolic blood pressure were taken and their means were used in the analysis. Total cholesterol levels were measured using automated enzymatic colorimetric methods.

6.3.4 Descriptive analysis

Baseline characteristics were investigated separately for individuals with and without fatal CHD or non-fatal MI and with and without fatal or non-fatal stroke during follow-up. Furthermore, characteristics of excluded versus included participants were compared. Differences between groups were assessed using the t-test and chi-squared test. A two-sided p-value <0.05 was considered statistically significant. To illustrate the variation of psychological distress scores within and across participants over time, individual psychological distress trajectories of 150 randomly selected participants with fatal CHD or non-fatal MI during follow-up were compared to individual psychological distress trajectories of 150 randomly selected individuals without events. This step was repeated for a random subset of 150 individuals each who did and did not have a fatal or non-fatal stroke by the end of follow-up.

6.3.5 Retrospective analysis of psychological distress trajectories

The aim of this analysis was to compare psychological distress trajectories of individuals prior to diagnosis of cardiovascular events to psychological distress trajectories of individuals free from cardiovascular events over the same time period. A latent process mixed model with psychological distress as dependent variable was fitted (see below). A mixed effects specification was used since repeated measures of the same participants are likely to be correlated. Time was modelled using a backward timescale starting at the date of diagnosis for individuals with events and the last date of follow-up or death among individuals without events (year 0). Time

dependence was modelled using natural cubic splines of time with knots set at the quartiles of the distribution. The event indicator was added to the model to allow for different intercepts among those with and without events, and an interaction term between time and the event indicator was added to allow for differences in the slope of the psychological distress trajectories. Furthermore, a random intercept and slope of time was added to the model to account for individual variation around the mean trajectory. In keeping with steps described in previous chapters (see section 5.3.3.2.2 Link function), the optimal link function to describe the relationship between the dependent variable (scores of psychological distress) and the linear predictor was identified by running models using five different link functions.

```
kn <- quantile(sample$time, probs = c(0.25, 0.50, 0.75))     #spline knots at the quartiles of the
                                                             #distribution
lcmm(fixed = GHQ ~ event*ns(time, knots = kn),
random = ~ time,
subject = "ID",
link = "linear",          #five different models using different link functions were run: linear,
beta,                     #equidistant I-splines, I-splines set at quantiles of distribution,
                          #threshold
data = sample)
```

Using the model with the preferred link function, mean predicted trajectories were plotted separately for those with and without events for each of the outcomes. Since differences between the trajectories of those with and without events might be due to differences in participant characteristics, analyses were subsequently adjusted for age at year 0 (year of diagnosis, death, or end-of follow-up), sex, and study phase. The Wald test was used to investigate if the addition of covariates and the interaction term between the event indicator and change of psychological distress over time improved model fit. A two-sided p-value <0.05 was considered statistically significant. The R package lcmm package was used to fit all models (Proust-Lima et al, 2017).

6.3.6 Group-based trajectory modelling among those with an event

Since the group of participants with cardiovascular events during follow-up might be composed of heterogeneous subgroups that differ in their trajectories of psychological distress over time, I investigated whether psychological distress subgroups (latent classes) could be identified. The steps involved in identifying potential latent classes were the same as those described as part of the group-based trajectory modelling of the ALSWH analysis (see section 5.3.3.2 Group-based trajectory modelling of depressive symptoms (CES-D)). In short, using the model with the preferred link function, the optimal number of classes was determined using statistical criteria with a prior requirement of at least 1% of all individuals in each of the latent classes. If the optimal number of classes exceeded one class, the optimal model structure would be determined by estimating models using model structures with increasing complexity (Table 5.1), and model adequacy and model discrimination would be investigated for each of the models that reached convergence in up to 500 iterations. The effect of confounders was assumed to be the same for all latent classes. The R code for a model with a two-class solution is presented below as an example.

```
lcmm(fixed = GHQ ~ ns(time, knots = kn) + AGE0 + SEX + WAVE,
mixture = ~ ns(time, knots = kn),
random = ~ time,
subject = "ID",
maxiter = 500,
ng = 2,
link = "linear",              #determined based on comparison of different link functions
data = sample)
```

6.4 Results

6.4.1 Study population

Out of 10,308 participants, 10,187 participants were included in the fatal CHD or non-fatal MI analysis and 10,252 participants were included in the analysis on fatal or non-fatal stroke (Figure 6.2).

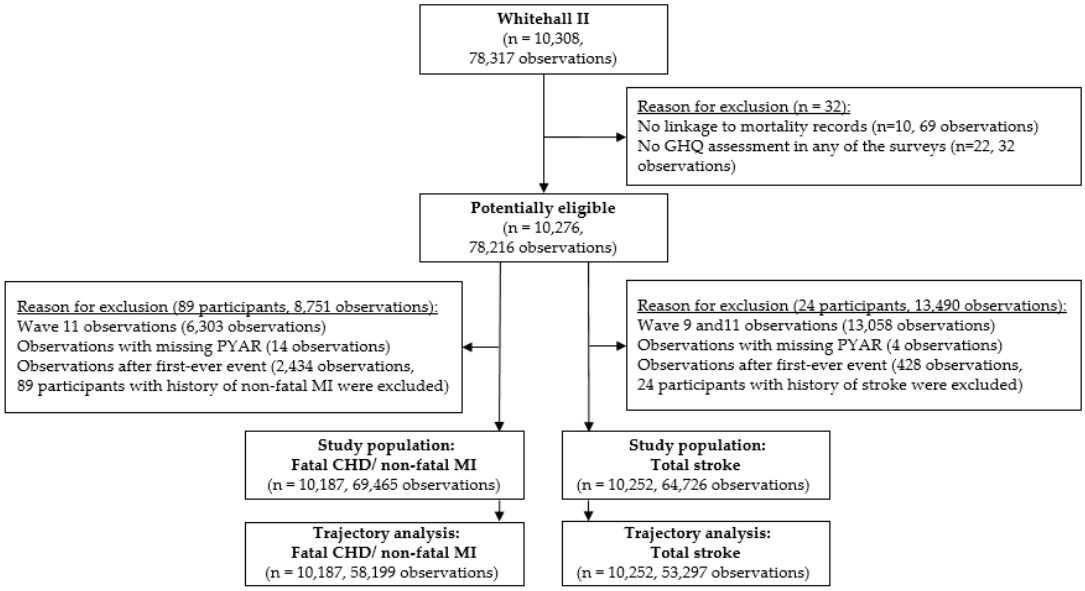


Figure 6.2: Flow chart of Whitehall II study participants included in analyses

There were statistically significant differences between individuals included and excluded from the analyses (Appendix Table A 40). Participants excluded from the fatal CHD or non-fatal MI analysis were more likely to be male, older, non-white, and clerical or support worker, and generally had worse cardiovascular risk profiles. Characteristics were similar between those included and excluded from analyses on fatal or non-fatal stroke with the exception that the proportion of male participants was higher among those included in the analyses (Appendix Table A 41).

6.4.2 Descriptive statistics

The median age (IQR) of participants included in the analysis on fatal CHD or non-fatal MI was 44.2 (39.6 – 50.3) years, the majority identified themselves as white (89.2%), and 6,803 (66.8%) were male (Table 6.1). The majority of participants were married (73.8%), and 48.1% were from professional or executive grades, whereas 22.5% and 29.4% were from clerical and office support grades and administrative

grades, respectively. The characteristics were similar for participants included in the analysis on fatal or non-fatal stroke (Table 6.1). There were 655 events of fatal CHD or non-fatal MI and 193 fatal or non-fatal stroke during follow-up. The median age (IQR) at fatal CHD or non-fatal MI was 59.8 (54.3 – 66.4) years and 63.6 (57.2 – 69.2) years at fatal or non-fatal stroke.

*Table 6.1: Characteristics of Whitehall II participants included in the analyses on fatal coronary heart disease or non-fatal myocardial infarction and fatal or non-fatal stroke at study phase 1 (page 1 of 2)**

	Population for CHD/ MI outcome (n = 10,187)	**Population for stroke outcome** (n = 10,252)
Male (%)	6,803 (66.8)	6,869 (67.0)
Age (median [IQR])	44.2 [39.6, 50.3]	44.3 [39.6, 50.3]
Ethnicity (%)		
White	9,084 (89.2)	9,140 (89.2)
Non-white	1,011 (9.9)	1,020 (9.9)
Missing value	92 (0.9)	92 (0.9)
Marital status (%)		
Married/ cohabiting	7,523 (73.8)	7,571 (73.8)
Other	2,627 (25.8)	2,643 (25.8)
Missing value	37 (0.4)	38 (0.4)
Occupational position (%)		
Administrative	2,996 (29.4)	3,022 (29.5)
Professional/ executive	4,898 (48.1)	4,922 (48.0)
Clerical/ support	2,293 (22.5)	2,308 (22.5)
Smoking status (%)		
Never smoker	5,023 (49.3)	5,044 (49.2)
Ex-smoker	3,232 (31.7)	3,258 (31.8)
Current smoker	1,861 (18.3)	1,878 (18.3)
Missing value	71 (0.7)	72 (0.7)
Alcohol intake (%)		
Risky drinking	2,394 (23.5)	2,412 (23.5)
Safe drinking	7,701 (75.6)	7,747 (75.6)
Missing value	92 (0.9)	93 (0.9)
Physical activity (median [IQR])	3.0 [1.0, 5.0]	3.0 [1.0, 5.0]
Fruit and vegetable intake (%)		
Less than once a day	4,240 (41.6)	4,267 (41.6)
Daily	4,231 (41.5)	4,255 (41.5)
More than daily	1,686 (16.6)	1,700 (16.6)
Missing value	30 (0.3)	30 (0.3)
Body mass index (kg/ m2) (median [IQR])	24.2 [22.3, 26.3]	24.2 [22.3, 26.4]
Diabetes (self-report or medication) (%)	95 (0.9)	98 (1.0)
Systolic blood pressure (median [IQR])	122.0 [113.0, 132.0]	122.0 [113.0, 132.0]

*Table 6.1 continued: Characteristics of Whitehall II participants included in the analyses on fatal coronary heart disease or non-fatal myocardial infarction and fatal or non-fatal stroke at study phase 1 (page 2 of 2)**

	Population for CHD/ MI outcome (n = 10,187)	**Population for stroke outcome** (n = 10,252)
Diastolic blood pressure (median [IQR])	76.0 [70.0, 83.0]	76.0 [70.0, 83.0]
Total cholesterol levels (median [IQR])	5.9 [5.2, 6.7]	5.9 [5.2, 6.7]
Psychological distress (median [IQR])	1.0 [0.0, 5.0]	1.0 [0.0, 5.0]
Psychological distress (binary) (%)		
No	7,372 (72.4)	7,419 (72.4)
Yes	2,721 (26.7)	2,737 (26.7)
Missing value	94 (0.9)	96 (0.9)
Time to event, death, or end of follow-up (mean (SD))	21.3 (8.0)	19.1 (6.4)
Event (whole follow-up) (%)	655 (6.4)	193 (1.9)
Age at diagnosis/ end of follow-up (median [IQR])	66.8 [61.5, 72.8]	64.0 [59.5, 70.2]

* Data are given as n (%) unless specified

CHD: coronary heart disease, IQR: interquartile range, MI: myocardial infarction, SD: standard deviation

There were statistically significant baseline differences between participants with and without events during follow-up (Table 6.2, Appendix Table A 42). Fatal CHD or non-fatal MI and fatal or non-fatal stroke during follow-up was associated with older age, non-white ethnicity, current smoking status, higher BMI, and higher systolic and diastolic blood pressure at baseline. The proportion of male participants was higher among individuals with an event but the sex difference was not statistically significant between individuals with and without stroke during follow-up. Furthermore, there were statistically significant differences between individuals with and without fatal CHD or non-fatal MI with respect to marital status, fruit and vegetable intake, and total cholesterol levels but not between individuals with and without stroke during follow-up. No differences between median psychological distress scores were observed for each of the outcomes but the proportion of individuals with high psychological distress was higher among the groups with events.

*Table 6.2: Characteristics of Whitehall II participants at study phase 1, separately for those with and without fatal coronary heart disease or non-fatal myocardial infarction during follow-up**

	No CHD/ MI (n = 9,532, 93.6%)	**CHD/ MI** (n = 655, 6.4%)	p
Male (%)	6,285 (65.9)	518 (79.1)	<0.001
Age (median [IQR])	44.0 [39.5, 50.0]	48.4 [42.5, 52.8]	<0.001
Ethnicity (%)			<0.001
White	8,528 (89.5)	556 (84.9)	
Non-white	916 (9.6)	95 (14.5)	
Missing value	88 (0.9)	4 (0.6)	
Marital status (%)			0.031
Married/ cohabiting	7,031 (73.8)	492 (75.1)	
Other	2,470 (25.9)	157 (24.0)	
Missing value	31 (0.3)	6 (0.9)	
Occupational position (%)			0.432
Administrative	2,808 (29.5)	188 (28.7)	
Professional/ executive	4,568 (47.9)	330 (50.4)	
Clerical/ support	2,156 (22.6)	137 (20.9)	
Smoking status (%)			<0.001
Never smoker	4,744 (49.8)	279 (42.6)	
Ex-smoker	3,028 (31.8)	204 (31.1)	
Current smoker	1,694 (17.8)	167 (25.5)	
Missing value	66 (0.7)	5 (0.8)	
Alcohol intake (%)			0.629
Risky drinking	2,230 (23.4)	164 (25.0)	
Safe drinking	7,216 (75.7)	485 (74.0)	
Missing value	86 (0.9)	6 (0.9)	
Physical activity (median [IQR])	3.0 [1.0, 5.0]	3.0 [1.0, 6.0]	0.271
Fruit and vegetable intake (%)			0.022
Less than once a day	3,949 (41.4)	291 (44.4)	
Daily	3,963 (41.6)	268 (40.9)	
More than daily	1,595 (16.7)	91 (13.9)	
Missing value	25 (0.3)	5 (0.8)	
Body mass index (kg/ m2) (median [IQR])	24.1 [22.3, 26.3]	25.2 [23.4, 27.2]	<0.001
Diabetes (self-report or medication) (%)	87 (0.9)	8 (1.2)	0.559
Systolic blood pressure (median [IQR])	122.0 [113.0, 132.0]	127.0 [117.0, 139.0]	<0.001
Diastolic blood pressure (median [IQR])	76.0 [70.0, 83.0]	80.0 [74.0, 88.0]	<0.001
Total cholesterol levels (median [IQR])	5.8 [5.1, 6.6]	6.4 [5.7, 7.2]	<0.001
Psychological distress (median [IQR])	1.0 [0.0, 5.0]	1.0 [0.0, 6.0]	0.421
Psychological distress (binary) (%)			0.027
No	6,903 (72.4)	469 (71.6)	
Yes	2,535 (26.6)	186 (28.4)	
Missing value	94 (1.0)	0 (0.0)	
Time to event, death, or end of follow-up (mean (SD))	21.9 (7.8)	12.7 (6.9)	<0.001
Age at diagnosis/ end of follow-up (median [IQR])	67.2 [62.2, 73.1]	59.8 [54.3, 66.4]	<0.001

* Data are given as n (%) unless specified, CHD: coronary heart disease, IQR: interquartile range, MI: myocardial infarction, SD: standard deviation

Whilst there was variation between psychological distress trajectories of different participants and variation across measurements of the same individual over time, little difference was observed between psychological distress trajectories of those with and without events during follow-up (Figure 6.3 and Figure 6.4). Of note, the scarcity of data points at about 15 to 20 years before diagnosis or end of follow-up was data driven since psychological distress was not assessed in the fourth study phase. Also, there were no psychological distress measures exactly at year 0 (year of diagnosis, death, or end-of follow-up) since events were identified through linkage to medical records.

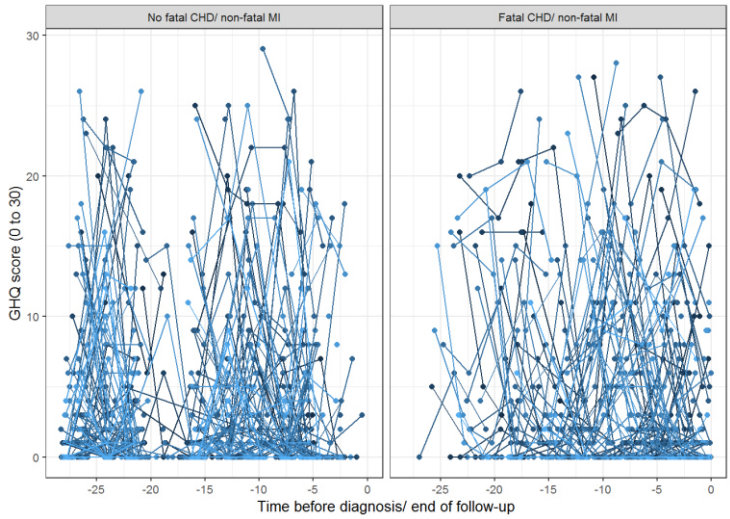


Figure 6.3: Psychological distress trajectories of 150 participants prior to diagnosis of fatal CHD or non-fatal myocardial infarction and 150 participants prior to death or end of follow-up

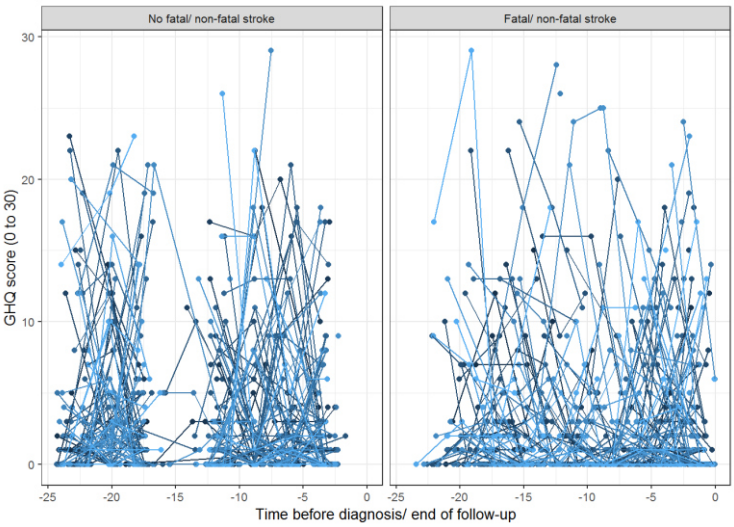


Figure 6.4: Psychological distress trajectories of 150 participants prior to diagnosis of fatal or non-fatal stroke and 150 participants prior to death or end of follow-up

6.4.3 Retrospective psychological distress trajectories of those with and without fatal CHD or non-fatal myocardial infarction

Five models with the same model structure but different link functions were fitted to identify the link function that provided the best model fit (as defined by the lowest discrete AIC). The approximation of the linear link function was relatively far from the approximations of non-linear link functions (Figure 6.5). Whilst the thresholds link function provided the lowest discrete AIC value, the model was computationally very intensive. Balancing statistical criteria of model fit and computational complexity, the splines link function with knots set at the quantiles of the distribution was chosen as preferred link function since it had the second lowest discrete AIC value and was computationally much less intensive (Table 6.3).

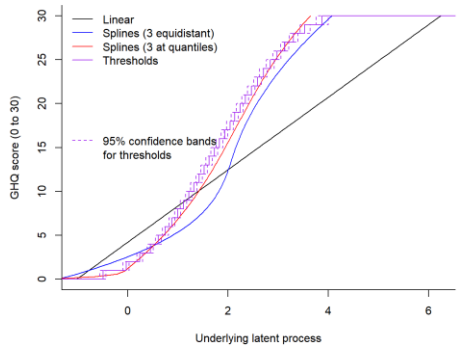


Figure 6.5: Estimated link functions of psychological distress (GHQ, range 0 to 30) in a latent process mixed model using natural cubic splines of time before fatal CHD, non-fatal myocardial infarction, or end of follow-up

Table 6.3: Model fit of five latent process mixed models using different link functions and natural cubic splines of time before fatal CHD, non-fatal myocardial infarction, or end of follow-up

Link function	Number of iterations	Number of parameters	Discrete AIC	Δ AIC*
Linear	42	14	252,629	25,785
Beta	500	Model did not reach convergence		---
Splines (equidistant)	25	17	243,135	16,290
Splines (at quantiles)	37	17	228,196	1,352
Thresholds	14	42	226,844	---

* Difference in AIC between model using thresholds link function and models using other link functions

In a model with the preferred link function, the unadjusted mean predicted psychological distress trajectories were estimated separately for individuals with and without fatal CHD or non-fatal MI (Figure 6.6). The mean predicted psychological distress trajectory of individuals with a fatal CHD or non-fatal MI event differed from the mean psychological distress trajectory of individuals without events. Throughout follow-up psychological distress scores were higher for people that had an event than those that did not. However, differences in mean scores were small and the CI of the group with a fatal CHD or non-fatal MI event was wide and overlapped with that for

people without events around year 0 and over 23 years before the end of follow-up. This was likely due to small numbers of individuals with measurements of psychological distress at these time points. The between-group difference in mean scores of psychological distress was relatively stable over time.

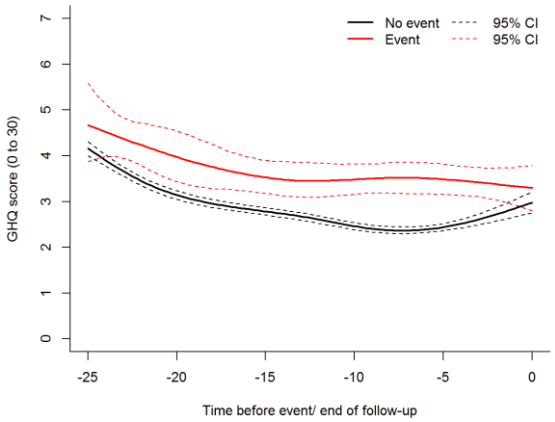


Figure 6.6: Unadjusted mean predicted psychological distress trajectories, separately for those with and without fatal coronary heart disease or non-fatal myocardial infarction

Since differences between psychological distress trajectories among those with and without events might be due to differences in participant characteristics, the analysis was subsequently adjusted for age at year 0 (year of diagnosis, death, or end-of follow-up), sex, and study phase. Whilst the addition of age at year 0 and sex improved model fit ($p < 0.01$, and $p < 0.01$, respectively), there was no statistically significant improvement of model fit due to the addition of study phase ($p = 0.20$). Furthermore, the interaction term between the event indicator and cubic splines of time did not improve model fit in either the unadjusted or adjusted model ($p = 0.08$ and $p = 0.14$, respectively). This indicated that there was no statistically significant difference in change of psychological distress scores over time between individuals with and without events. As a result, mean predicted trajectories of psychological distress were modelled with an adjustment for the event indicator on the intercept only whilst additionally adjusting for age at year 0 and sex (Figure 6.7). The adjustment for the event indicator on the intercept provided statistically significant improvement of model fit ($p < 0.01$) indicating that there is a statistically significant difference in mean psychological distress scores between groups. Of note, due to the

exclusion of the interaction term this between-group difference had to be constant over time. Among both men and women, the mean predicted scores of psychological distress were slightly higher among those with a fatal CHD or non-fatal MI than among those without an event. Furthermore, the mean predicted score of psychological distress was very similar for women without events and men with events. A formal test of interaction between sex and change of psychological distress scores over time among those with and without events was not performed since numbers of events were considered too small to add a three-way interaction term.

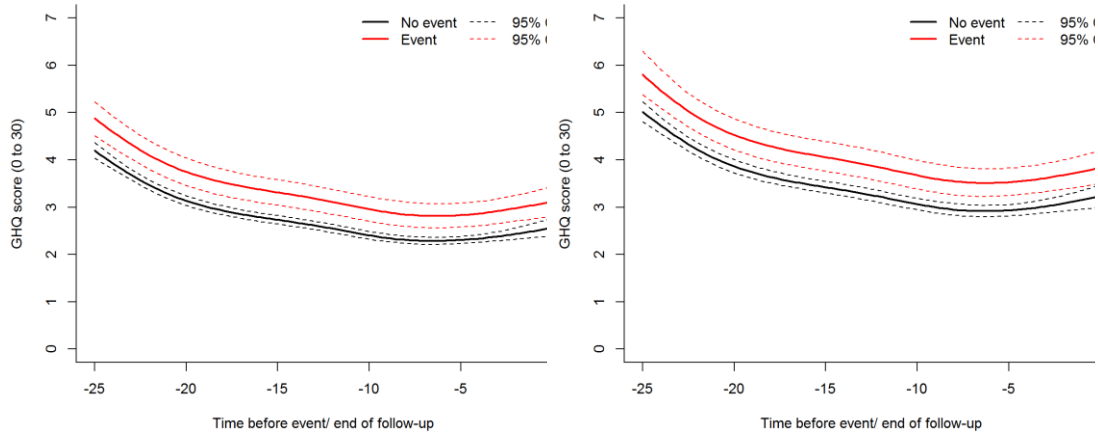


Figure 6.7: Mean predicted psychological distress trajectories of a hypothetical man aged 65 years at year 0 (left) and a hypothetical woman aged 65 years at year 0, separately for those with and without fatal coronary heart disease or non-fatal myocardial infarction

6.4.4 Group-based trajectory modelling among those with fatal CHD/ non-fatal MI

There was no evidence that the group of individuals with fatal CHD or non-fatal MI was composed of multiple latent classes. Using the preferred link function identified in the previous analysis, models with increasing numbers of latent classes were fitted until two models did not reach convergence in up to 500 iterations (Table 6.4). The discrete AIC values of all models were very close. The lowest discrete AIC value was observed in the model with a one class solution. Interestingly, the model with a three class solution offered the second lowest discrete AIC value and lowest BIC value. An investigation of the mean predicted trajectories of all models that reached convergence showed that the three class solution closely approximated the three class solution that was identified in the ALSWH analysis (Appendix Figure A 5). However,

since the one class solution had the lowest discrete AIC value, no further analysis was performed.

Table 6.4: Model characteristics of models with increasing numbers of latent classes among those with fatal coronary heart disease or non-fatal myocardial infarction

Number of groups	Number of iterations	Log-likelihood	Number of parameters	Discrete AIC	BIC	% class 1	% class 2	% class 3	% class 4
1	26	-5,192	14	11,013	10,475	100.0			
2	131	-5,161	20	11,027	10,451	19.1	80.9		
3	61	-5,140	26	11,022	10,448	25.3	4.0	70.7	
4	48	-5,140	32	11,034	10,487	28.1	4.6	67.3	0
5	500	Model did not reach convergence							
6	500	Model did not reach convergence							

6.4.5 Retrospective psychological distress trajectories of those with and without fatal or non-fatal stroke

The comparison of the approximations of five different link functions provided very similar results using cubic splines of time before fatal or non-fatal stroke or end of follow-up (Figure 6.8). Again, the threshold link function provided the lowest discrete AIC value. However, since it was computationally intensive the splines link function with knots set at the quantiles of the distribution was chosen balancing criteria of model fit and computational complexity (Table 6.5).

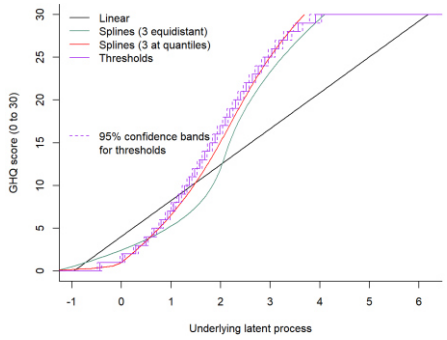


Figure 6.8: Estimated link functions of psychological distress (GHQ, range 0 to 30) in a latent process mixed model using natural cubic splines of time before fatal or non-fatal stroke or end of follow-up

Table 6.5: Model fit of five latent process mixed models using different link functions and natural cubic splines of time before fatal or non-fatal stroke or end of follow-up

Link function	Number of iterations	Number of parameters	Discrete AIC	Δ AIC*
Linear	17	14	235,988	23,161
Beta	500	Model did not reach convergence		---
Splines (equidistant)	76	17	227,323	14,496
Splines (at quantiles)	250	17	214,036	1,209
Thresholds	56	42	212,827	---

* Difference in AIC between model using thresholds link function and models using other link functions

There was little evidence for differences in mean predicted psychological distress scores among those with and without fatal or non-fatal stroke in unadjusted analysis (Figure 6.9). The mean predicted psychological distress scores of the group with events were slightly higher than those of the group without events from year 0 up to around 18 years prior to diagnosis, death or end of follow-up. In years 19 and 20 before diagnosis, death or end of follow-up, the mean score of the group without events was slightly higher than that of individuals with events. Furthermore, there seemed to be slightly more variation in psychological distress scores among those with events over time. However, due to small numbers CIs of the group with events were wide, especially in the years 17 to 20 before diagnosis.

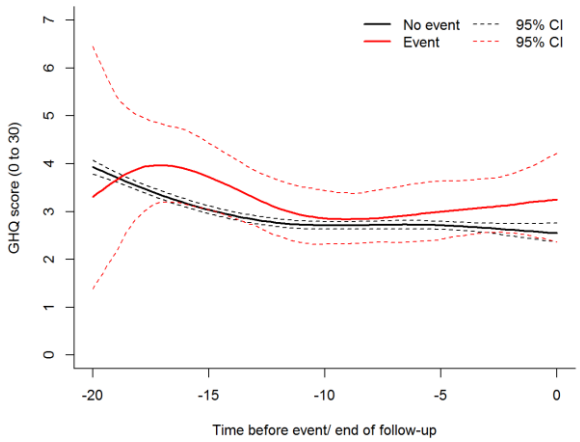


Figure 6.9: Unadjusted mean predicted psychological distress trajectories, separately for those with and without fatal or non-fatal stroke

There was some evidence for differences in mean predicted psychological distress scores among those with and without fatal or non-fatal stroke in adjusted analysis. Whilst adjustments for age at year 0 and sex improved model fit ($p < 0.001$, and $p < 0.001$, respectively), adjustment for study phase and the interaction term between the event indicator and change of psychological distress scores over time did not ($p = 0.12$, and $p = 0.44$, respectively). As described previously, this suggested that there was no statistically significant difference between those with and without events with regard to change of psychological distress scores over time. Therefore, models were subsequently fitted with an adjustment for the event indicator on the intercept of the psychological distress trajectories only whilst simultaneously adjusting for age at

year 0 and sex. The mean predicted psychological distress scores of a hypothetical man and woman aged 65 years at the time of diagnosis, death or end-of follow-up slightly decreased over time (Figure 6.10). As observed with the mean predicted scores of psychological distress prior to diagnosis with fatal CHD and non-fatal MI, the mean predicted scores of psychological distress were higher among women than men and higher among individuals with events than among those without events. In contrast to the analysis on psychological distress trajectories among those with or without fatal CHD or non-fatal MI, the CIs of the groups with and without fatal or non-fatal stroke overlapped. In keeping with that the adjustment for the event indicator on the intercept of the psychological distress trajectories did not provide statistically significant improvements of model fit ($p = 0.17$) suggesting that there is no statistically significant difference in mean predicted psychological distress scores between those with and without events. As described previously, the patterns of psychological distress over time had to be the same among those with and without events in this model since the interaction term between the event indicator and change of psychological distress over time was not included in the model. Following the same decision as described previously, it was not investigated if there was an interaction between sex and change of psychological distress over time among those with and without events.

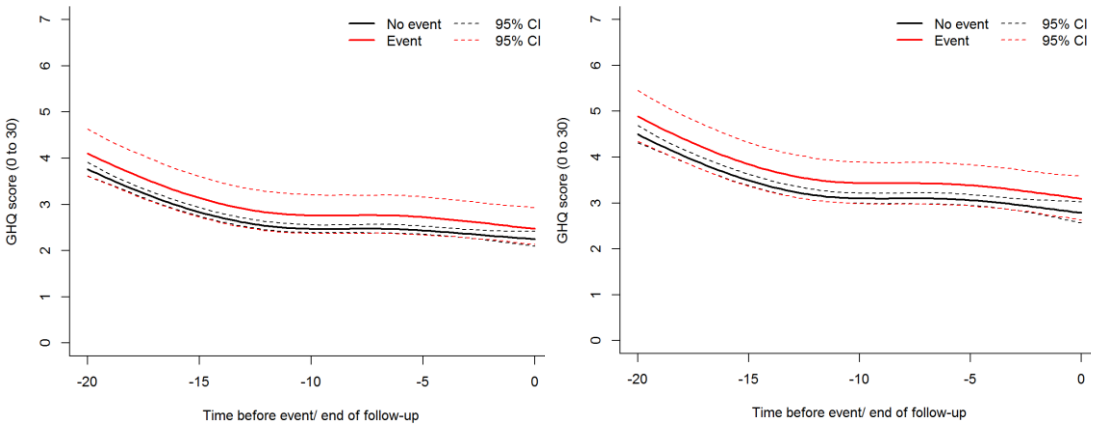


Figure 6.10: Mean predicted psychological distress trajectories of a hypothetical men aged 65 years at year 0 (left) and a hypothetical women aged 65 years at year 0 (right), separately for those with and without fatal or non-fatal stroke

6.4.6 Group-based trajectory modelling among those with fatal or non-fatal stroke

In keeping with results among those with fatal CHD or non-fatal MI, there was no evidence that the group of individuals with fatal or non-fatal stroke was composed of multiple latent classes. Models with one and two latent classes reached convergence in less than 500 iterations, whereas models with three and four classes did not reach convergence. The model with one latent class had the lowest discrete AIC indicating best model fit (Table 6.6). The mean predicted trajectories of psychological distress for a hypothetical man and women aged 65 years at year 0 are presented in Appendix Figure A 6.

Table 6.6: Model characteristics of models with increasing number of latent classes among those with fatal or non-fatal stroke

Number of groups	Number of iterations	Log-likelihood	Number of parameters	Discrete AIC	BIC	% class 1	% class 2
1	22	-1,634	14	3,548	3,341	100	
2	58	-1,621	20	3,557	3,347	13.0	87.0
3	500	Model did not reach convergence					
4	500	Model did not reach convergence					

6.5 Discussion

6.5.1 Summary of main findings

There was some evidence of differences between mean predicted psychological distress scores among those with and without cardiovascular events in the time prior to diagnosis, death, or end-of follow-up but no clear evidence for differences in the patterns of psychological distress over time. The mean predicted scores of psychological distress were higher among those with fatal CHD or non-fatal MI than among those who did not have an event over the same time period. The difference in psychological distress scores between groups was constant over time. Although CIs did not overlap, the difference between groups was small and might not be clinically meaningful. A similar pattern was observed when psychological distress trajectories among those with fatal or non-fatal stroke were compared to those who remained free from stroke over the same time period. However, the difference between groups was even smaller, CIs overlapped, and adjusting for the event indicator did not

suggest a statistically significant difference between mean scores of psychological distress. For both outcomes, the intercepts of the psychological distress trajectories were higher among women than men. Differences in the shape of the trajectories among men and women with and without events were not investigated since numbers of events were considered too small to add a three-way interaction term.

There was no evidence for latent classes with varying psychological distress trajectories among individuals with fatal CHD or non-fatal MI and those with fatal or non-fatal stroke during follow-up. Although trajectories of psychological distress differed between individuals in descriptive analyses, the one-class solution offered the lowest discrete AIC value for both outcomes. Of note, the discrete AIC values of models with varying number of classes were very close to each other indicating that they provide good alternatives to each other. Interestingly, the three class solution of psychological distress trajectories among those with a fatal CHD or non-fatal MI was very similar to the three psychological distress trajectories identified as part of the ALSWH analysis that suggested stable low, stable moderate and a fluctuating pattern of depressive symptoms over time. The model with a three class solution among those with fatal or non-fatal stroke did not reach convergence in less than 500 iterations, possibly due to lack of statistical power as a result of the lower number of individuals with fatal or non-fatal stroke during follow-up.

6.5.2 Strength and weaknesses of study

6.5.2.1 Chance

The data requirements with regards to sample size and number of data points (see section 5.3.3.2.1 Latent process mixed modelling) were met in this analysis. However, lack of statistical power might have influenced findings, especially in the analysis on fatal or non-fatal stroke. First, p-values were used to identify the probability of finding the observed or more extreme baseline differences between those with and without events during follow-up given that the null hypothesis of no difference between groups was true. Baseline differences were considered statistically significant at the 5% significance level. Lack of statistical significance might indicate

that there truly was no difference between groups or that the analysis lacked statistical power to identify differences between groups. This might have been particularly important among those with and without stroke during follow-up due to the low number of events. Furthermore, p-values were used to detect statistically significant improvements of model fit using the Wald test. For both outcomes neither the interaction term between change of psychological distress score over time and the event indicator nor the covariate study phase significantly improved the model fit at the 5% significance level. Again, this might be because there truly was no interaction and improvement of model fit. However, a different explanation could be lack of statistical power to detect an interaction. Second, the possibility of chance was considered by providing CIs. Generally, bigger sample sizes and larger number of events improve the precision of estimates and narrow CIs. CIs were relatively narrow in the retrospective trajectory analyses among those with and without fatal CHD or non-fatal MI and among those with and without fatal or non-fatal stroke. However, CIs of psychological distress trajectories did not overlap in the analysis on fatal CHD or non-fatal MI whereas they did overlap between those with and without stroke during follow-up. Overlapping CIs indicate that differences between groups might have been due to chance alone. As described above, one potential explanation for this finding was the smaller number of events in the analysis on fatal or non-fatal stroke. In order to investigate this further, it would be of interest to repeat the analysis once the information on events in study phases 9 and 11 were validated by the Whitehall II study team. In addition, CIs of trajectories of potential latent classes were wide and overlapped. The one class solution was the preferred solution in this analysis. However, it has been discussed that one disadvantage of latent class analysis is the potential lack of statistical power due to differently sized subgroups and small numbers among latent classes (Vistisen et al, 2014). Therefore, the one class solution might have been identified as the preferred option because there was insufficient power to identify other classes.

6.5.2.2 Bias

This analysis has major advantages over existing studies. One major strength of the analytical approach is the use of repeated measures of psychological distress. The use of repeated measures over time allowed for fluctuating exposure levels across measurements of the same individual. This is a major advantage over existing studies since those frequently relied on the use of a baseline measure of the exposure (see section 3.4.2.3 Exposure assessment for more detail). Furthermore, due to the use of scores of psychological distress the creation of an arbitrary cut-off was avoided. Applying cut-offs to identify individuals with and without psychological distress might lead to misclassification of individuals, especially misclassification of those with scores close to the cut-off value. A further strength is the long follow-up that allowed an assessment of psychological distress trajectories over around 25 years. The long follow-up was particularly important since CVD are often preceded by a long subclinical phase of atherosclerosis development. For example, one study among individuals aged 65 years and older suggested that 62% of individuals without clinical CVD had subclinical CVD (Kuller et al, 2006). Since the prevalence of subclinical CVD increases with age, younger participants are less likely to have subclinical CVD at baseline. Participants of the Whitehall II study were aged between 35 to 55 years at baseline. Ideally, participants would have been even younger at baseline but the advantages of the Whitehall II study with regards to the rich information on psychological distress over time outweighed its disadvantage of not having an even younger sample at baseline.

One disadvantage of the analysis is that the GHQ scale is a non-specific measure of psychological distress rather than a measure of depressive symptoms. The GHQ and CES-D rating scales have been validated against the revised Clinical Interview Schedule (CIS-R), a structured diagnostic interview for common mental disorders (Lewis et al, 1992), using a random sample of Whitehall II participants supplemented with Whitehall II participants with depression (Head et al, 2013). Sensitivity and specificity values for detecting any mental disorder were 86% and 87% for the GHQ and 77% and 89% for the CES-D. Sensitivity and specificity for detecting depressive

episode were 78% and 83% for the GHQ and 89% and 86% for the CES-D. Head et al (2013) concluded that both scales had good validity for detecting any mental disorders and depressive episodes. However, in the investigation of the validity of the GHQ scale for detecting depressive episodes a four-item subscale was identified from the 30-item GHQ scale which was not used in this analysis. Therefore, it was likely that the GHQ scores used in this analysis represented a measure of non-specific psychological distress rather than a measure of depressive symptoms. It would have been interesting to compare results using the psychological distress rating scale with results using the CES-D scale. Unfortunately, the CES-D scale was first introduced in study phase 7. Whilst there were repeat assessments of depressive symptoms in study phases 9 and 11, these study phases could not be used in this analysis due to the lack of validated outcomes. Once outcomes are validated for these waves and additional assessments of depressive symptoms are collected, it will be of interest to compare the findings based on psychological distress with findings based on depressive symptoms. However, it should be noted that this would not have influenced the internal validity of results but reduced the comparability of findings with other parts of this project.

The Whitehall II study population is not representative of the general population. The characteristics of the study population reflect the composition of the Civil Service at study baseline (Marmot & Brunner, 2005). Most participants are white collar workers in stable employment, there are no manual worker in the sample, and two thirds of the sample are men. As a result, the prevalence of high psychological distress might be lower in the Whitehall II study population than in the general population. Furthermore, the target population only included civil servants in London offices thereby limiting the geographical generalisability.

Loss to follow-up might have introduced selection bias. In study phase 8 (last phase used in analyses on fatal or non-fatal stroke) 75.2% of those alive participated whereas 72.3% of those alive participated in study phase 9 (last phase of analyses on fatal CHD/ non-fatal MI) (University College London, 2019). Since information of those

lost to follow-up was missing, the analysis was inevitably restricted to those not lost to follow-up. Since those not lost to follow-up might have differed in their participant characteristics, the analysis might be influenced by selection bias. Additionally, the selection of participants might have influenced the generalisability of findings. Participants of the Whitehall II study were British civil servants of mainly white ethnicity. Whilst this would not have influenced the internal validity of findings, it might have influenced the external validity.

Psychological distress might have been underreported due to the stigma attached to mental disorders. If there was reporting bias, the mean scores of psychological distress would likely be underestimated. This effect would have been non-differential between individuals with and without cardiovascular events since these events had not happened at the time of reporting psychological distress. Furthermore, it seems unlikely that the effect of underreporting differed over time. As a result, underreporting would have biased findings towards the null. In this analysis, men had consistently lower mean score of psychological distress than women. If men were more likely to underreport depressive symptoms, the lower mean scores of psychological distress might reflect reporting bias rather than true differences in mean scores among men and women.

6.5.2.3 Confounding

Confounding might have influenced findings. Results were adjusted for sex, age at year 0 and study phase, all of which were potential time-constant confounding factors. Whilst no adjustment is perfect, it is likely that it accounted for most of difference in age profiles in the groups with and without events. However, it should also be noted that there was longer follow-up of those without events than those with events. It is likely that this affected the precision of the estimation of trajectories of those with events at longer follow-up.

An existing publication using the Whitehall II study that investigated the trajectories of psychological distress and depressive symptoms before diagnosis of dementia additionally adjusted their analysis for ethnicity and education. I decided not to

adjust for ethnicity in this analysis since the vast majority of Whitehall II participants were of white ethnicity. Furthermore, the analysis was not adjusted for education since that information was not available at baseline. Furthermore, there were no baseline differences in occupational position between those with and without events during follow-up. Additionally, it was decided not to adjust the analysis for time-varying factors since events during follow-up might influence the trajectories as potential mediating factors. It would have been of interest to investigate if there were differences in the shape of psychological distress trajectories over time among men and women with and without events. This was not investigated since numbers of events were considered too small to add a three-way interaction term.

6.5.3 Comparison of findings with previous research

This is the first study that compared psychological distress trajectories prior to diagnosis with cardiovascular events to psychological distress trajectories among individuals free from cardiovascular events over the same time period. Thus, this study complements existing studies by investigating if reverse causation or overlap of somatic symptoms might explain the association between psychological distress and CVD. Due to the use of a retrospective timescale, an advantage over existing studies is that this study did not create a window in which the exposure was defined and assessed the risk of outcomes in the period after the exposure window but used information on the exposure in all waves prior to cardiovascular events, death, or end of follow-up. Also, the use of psychological distress as continuous measure allowed for an assessment of the change in psychological distress over time without creating an arbitrary cut-off. Whilst it might be argued that a comparison of those with high and low psychological distress scores is clinically more meaningful, the use of continuous scales is advantageous from a statistical power standpoint.

Three existing studies have investigated the relationship between psychological distress, depressive symptoms and risk of CHD and stroke using data from the Whitehall II study (Brunner et al, 2014; Nabi et al, 2008; Stansfeld et al, 2002). Stansfeld et al (2002) reported that the odds of self-reported CHD at five years of follow-up

were higher among men and women with high psychological distress, compared to those with low psychological distress at baseline. In keeping with that, Brunner et al (2014) reported higher hazards of fatal CHD or non-fatal MI among individuals with high psychological distress, relative to those with low psychological distress over five years and over ten years of follow-up. A five year lag period was used in the latter analysis to minimise the potential effect of reverse causation. Furthermore, high psychological distress was associated with higher hazards of fatal or non-fatal stroke over five years but not over ten years of follow-up using a five year lag period (Brunner et al, 2014). There was evidence of a dose-response relationship between number of study phases with high psychological distress and hazard of fatal CHD or non-fatal MI but not for risk of fatal or non-fatal stroke. However, estimates were imprecise due to low number of events. Additionally, there were higher hazards of fatal CHD or non-fatal MI and fatal or non-fatal stroke among those with high depressive symptoms, relative to those with low depressive symptoms over 2.5 years of follow-up. Brunner et al (2014) concluded that there was evidence of an association and dose-response relationship between psychological distress and fatal CHD or non-fatal MI whereas the association between psychological distress and fatal or non-fatal stroke was explained by reverse causation. The observed dose-response effect was only partially seen in a different study using data from the Whitehall II study (Nabi et al, 2008). Using information from two study phases to identify individuals with high psychological distress in none, one, and two of the study phases, there were increased hazards of fatal CHD or non-fatal MI among individuals with high psychological distress in two study phases but not in one study phase, relative to those with low psychological distress in all study phases over a mean follow-up of 11.1 years. Differences between the study results might be explained by differences in the categorisation of the cumulative effect of psychological distress. Whilst Nabi et al (2008) used three categories of none of the times, once, and two times to investigate the effect of cumulative phases of high psychological distress on fatal CHD or non-fatal MI, Brunner et al (2014) used categories of none of the times, one to two times, and three to four times to define the exposure.

The pattern of the trajectories of psychological distress prior to diagnosis of cardiovascular events was different from the pattern of psychological distress before diagnosis of dementia. Using data from the Whitehall II study, Singh-Manoux et al (2017) observed a steep increase in psychological distress scores a decade before diagnosis of dementia whilst scores of psychological distress remained low among those free from dementia over the same time period. In contrast, this analysis showed slightly higher mean scores among those with cardiovascular events than among those without cardiovascular events throughout follow-up but the difference between mean scores of psychological distress was constant over time.

Whilst the results by Singh-Manoux et al (2017) seemed to support that high psychological distress prior to diagnosis of dementia was a prodromal feature of dementia itself, this was an unlikely explanation for observed differences between those with and without cardiovascular events in this analysis. One potential explanation why there might not have been clear evidence for differences between those with and without events in terms of the change of psychological distress over time might be lack of power to detect statistically significant interaction between the event indicator and change of psychological distress over time. However, this explanation seems unlikely, at least for the outcome of fatal CHD/ non-fatal MI, given that there was sufficient statistical power to detect interaction in the study by Singh-Manoux et al (2017). Singh-Manoux et al (2017) used the same data source and the number of participants with dementia during follow-up was lower than the number of participants with fatal CHD/ non-fatal MI in this analysis (n = 322 and n = 655, respectively). Furthermore, findings are unlikely explained due to chance alone since results were consistent for both fatal CHD or non-fatal MI and fatal or non-fatal stroke. The common soil hypothesis might explain why participants with events had consistently higher mean psychological distress scores before time of diagnosis, death or end of follow-up. As described previously some researchers believe that depression and CVD share common roots (common soil hypothesis), but the effect of these common roots manifests earlier on one disease than the other (see section 2.3.2.3 Proposed mechanisms). Examples of potential shared risk factors are shared genetic

variants (genetic pleiotropy) (de Geus, 2006), early-life stressors or low SES (Mortimer et al, 2016). However, it seems unlikely that these common roots would have the same effect on psychological distress scores over 25 years of follow-up. A further potential explanation is that consistently higher mean scores of psychological distress are indeed a risk factor for cardiovascular events later in life. However, it should be noted that the difference between the trajectories of psychological distress were very small and CIs overlapped between those with and without fatal or non-fatal stroke which might suggest that the difference is not clinically meaningful. Therefore, a third potential explanation is that differences between mean scores of psychological distress were observed due to chance alone.

6.6 Conclusion

In a prospective study of British civil servants, there was some evidence of differences between mean psychological distress scores among those with and without cardiovascular events in the time prior to diagnosis, death, or end-of follow-up but no clear evidence for differences in the patterns of psychological distress over time. Thus, it is unlikely that reverse causation or overlap of symptoms between subclinical CVD and somatic symptoms of psychological distress explain the association between psychological distress and CVD in this study.

Chapter 7: Discussion

7.1 Introduction

Depression and CVD both substantially contribute to the global burden of diseases. An association between depression and CVD onset and progression has been observed in numerous primary studies and existing reviews but depression has not been conclusively identified as independent risk factor for CVD. Several potential mechanisms have been proposed, some of which causally relate depression to CVD whereas others suggest that similar aetiological processes or methodological shortcomings of the previous studies were likely to be responsible for the observed association.

This project aimed to improve our understanding of the relationship between depression and subsequent CVD. A systematic review and meta-analysis was performed to identify shortcomings of existing studies that could potentially be overcome in this project. Quantitative analyses of three cohort studies and linkage to administrative datasets were used to further our understanding of the relationship. The analyses benefitted from strengths of each of the datasets and were able to address some of the limitations of the prior evidence base.

This discussion is subdivided into three parts. The first part outlines the contribution of this PhD to our understanding of the relationship between depression and subsequent CVD and relates them to methodological shortcomings identified in the systematic review. The second part of this discussion highlights strengths and limitations of this project as a whole. It also contains a discussion of potential alternative explanations for the observed association that were outside the scope of this project but that could be addressed in future research. This section will not repeat the discussion of strengths, limitations and potential sources of error of the systematic review and primary data analyses since these issues have been discussed in detail in the respective chapters. In the third part of this discussion potential implications for research and practice are presented.

7.2 Contribution of this project to our understanding of the relationship between depression and subsequent CVD

This project furthered our understanding of the relationship between depression and subsequent CVD by investigating the merit of potential non-causal explanations of the observed association. First, the residual confounding hypothesis was investigated using data from the UK Biobank. Depression, defined in different ways, was associated with increased hazards of MCVE in unadjusted, partially adjusted and fully adjusted models. Furthermore, there was the suggestion of a possible dose-response relationship. The UK Biobank collected a great breadth of variables, allowing for adjustment for a wide range of potential confounding factors. Despite some potential remaining residual confounding, the strength and consistency of the association between depression and hazard of MCVE strongly suggests that the association between depression and CVD is unlikely to be due to residual confounding alone. It should also be noted that some of the covariates adjusted for in the analysis may have occurred after the onset of depression and therefore may be on the causal pathway. Due to their genetic determination and the age of our sample participants, it is likely that the covariates of the partially adjusted model occurred before the onset of depressive symptoms. In contrast, some covariates of the fully adjusted model, such as weight gain or the onset of comorbidities, may have occurred after the onset of depression. If so, the partially adjusted model, which provided stronger effect estimates than the fully adjusted model, may have offered better estimates of the total effect of depression on MCVE risk.

Second, the hypothesis that the observed association was due to reverse causation or misclassification of individuals due to the overlap between somatic symptoms of depression and symptoms of subclinical CVD was investigated using data from the Whitehall II study. Previous studies tried to address the issue of reverse causation by excluding participants with events in the first months or years of follow-up (Brunner et al, 2014; Everson et al, 1998; Gustad et al, 2014; Scherrer et al, 2011; Stürmer et al, 2006; Wouts et al, 2008), by using depressive rating scales that did not contain somatic symptoms, such as the HADS-D or GDS (Gustad et al, 2014; Köhler et al, 2013;

Langvik & Hjemdal, 2015), or by conducting sensitivity analyses in which the authors excluded items from the depressive symptom rating scale that captured somatic symptoms (Barefoot & Schroll, 1996). The analysis based on data from the Whitehall II study was the first to address the reverse causation hypothesis in a more sophisticated way using repeat assessments of psychological distress before diagnosis of cardiovascular events. There was some evidence of higher mean psychological distress scores among those with cardiovascular events than among those without cardiovascular events prior to diagnosis, death, or end of follow-up. However, the difference in mean scores of psychological distress was small and constant over time. If higher psychological distress scores were indeed an indication of prodromal CVD itself, one would expect an increase in psychological distress scores in the years immediately prior to development of CVD. Therefore, reverse causation or misclassification of individuals due to an overlap in somatic symptoms did not seem to explain the observed association in this dataset. Instead the results suggest that psychological distress and CVD share common roots (common soil hypothesis) or that psychological distress is causally related to CVD.

In addition to investigating these non-causal explanations, the review highlighted that the vast majority of existing studies relied on single measures of depression or depressive symptoms. Taking into account that depressive symptoms may not be static over time, I identified three latent classes with stable low, stable moderate, and fluctuating depressive symptoms in the young cohort of the ALSWH. Importantly, the three latent classes did not only differ in their pattern of depressive symptoms over time but also were characterised by differences in terms of their cardiovascular risk profiles at baseline and end of follow-up. It was particularly interesting that the patterns of change of BMI over time differed between women with stable low, stable moderate, and fluctuating depressive symptoms. Although the mean BMI of all groups increased over time, the increase was more rapid among those with stable moderate and fluctuating depressive symptoms than among those with stable low depressive symptoms. The difference in mean scores of BMI at 20 years of age was relatively small. This result is of particular interest given that BMI is one of the

potential mediators between depression and subsequent CVD. Since this project was one of the first to investigate the change of BMI over time among women with different patterns of depressive symptoms, it is of interest to replicate this analysis using data from other cohort studies. As a next step, there is a need to investigate the relevance of different patterns of change of depressive symptoms in terms of risk of subsequent CVD and the role of BMI as mediating factor using more advanced statistical analyses (see section 7.4 Implications for research for more detailed discussion). Furthermore, these results may be important with regard to potential interventions aiming at the prevention of weight gain, especially among individuals with stable moderate and fluctuating depressive symptoms (see section 7.5 Implications for practice).

Moreover, this project highlighted that the association between depression and CVD was stronger among individuals from low socioeconomic backgrounds. The factors contributing to the increased risk of CVD among this group of participants are likely to be multifactorial. First, individual level factors such as behaviours, attitudes and habits might contribute to the observed association. For example, individuals with depression and a low socioeconomic status may be less likely to seek help than individuals with depression and a high socioeconomic status. This might be due to differences in the individual's attitudes towards health services or the availability of resources such as transport to reach services (Magaard et al, 2017; Roberts et al, 2018). The proportion of individuals that seek help might not align with the proportion of individuals that need treatment for depression. Given that individuals from low socioeconomic backgrounds are more likely to suffer from multimorbidity from early adulthood onwards (McLean et al, 2014), one would expect a higher need for services and treatment among these individuals. Furthermore, there was a large proportion of individuals with adverse health behaviours in this group of participants in the UK Biobank analysis. If less help-seeking occurs despite higher needs for services and treatment, the increased risk of CVD among individuals with depression and a low socioeconomic status may reflect the association between more severe or untreated depression and CVD. Second, factors of the living environment are likely to

contribute to the observed pattern. For example, factors related to health care services such as availability of services, waiting times and quality of care might affect the observed association. Using routine data from 956 general practices in Scotland McLean et al (2015) have investigated the relationship between multimorbidity, general practice resources, workload and deprivation. They concluded that general practices in more deprived areas carried out more consultations but the gradient did not align with the social gradient in need due to higher levels of multimorbidity and mortality. Furthermore, general practices in more deprived areas did not receive additional funding to meet the higher need in more deprived areas. These findings are of particular interest with regard to implications for practice (see section 7.5 Implications for practice).

Similarly, this project highlighted that the association between depression and CVD was stronger among those with physical comorbidities. If the comorbidities indeed occurred before the onset of depression, potential reasons for this observed pattern might be that depression might make self-management strategies more difficult among individuals with existing diabetes, hypertension or high cholesterol levels. For example, depression has been shown to be a risk factor for medical noncompliance among individuals with diabetes (DiMatteo et al, 2000; Gonzalez et al, 2008) which might lead to adverse effects such as poor glycaemic control among individuals with comorbid depression and diabetes (Lustman et al, 2000). It is also possible that individuals with mental-physical comorbidity receive worse quality of care which in turn might increase their risk of adverse events such as cardiovascular events (Jørgensen et al, 2017; Kurdyak et al, 2017). If instead the onset of depressive symptoms or the diagnosis of depression preceded the onset of diabetes, hypertension and high cholesterol levels, these physical comorbidities may be on the causal pathway between depression and CVD. Interestingly, individuals with depression but no hypertension were not at increased risk of CVD, relative to those with neither depression nor hypertension. However, the risk of individuals with depression and hypertension exceeded the sum of the risks due to depression alone and hypertension alone. Since this finding was not in keeping with previous research

(see section 4.5.3.4 The role of comorbidities), there is a need to replicate this analysis in future studies. Also, further research is needed to improve our understanding of the role of comorbidities in the relationship between depression and CVD (see section 7.4 Implications for research).

7.3 Strengths and limitations

7.3.1 The use of three cohort studies in this project

The use of three different datasets in this project offered advantages but also had disadvantages. One advantage of using three cohort studies and different analytical approaches was that various aspects of the relationship between depression and subsequent CVD were investigated. Thus, a number of shortcomings of the existing evidence base could be addressed. More specifically, I addressed potential explanations of the observed association such as residual confounding, reverse causation, and potential overlap of somatic symptoms of individuals with worsening vascular disease and depressive symptoms. Furthermore, I improved upon one of the main methodological shortcomings of the evidence base by using repeat assessments of depressive symptoms instead of a single measure and I identified factors that modified the relationship between depression and CVD. A further advantage of using three cohort studies and different analytical approaches was that it provided me with experience of using different datasets and novel methods. A disadvantage of using three datasets was that I could not use the same measure of depression and CVD for all analyses since different information was collected in each of the cohort studies. For example, whilst various measures of depression were used in analyses based on the UK Biobank, measures of depressive symptoms and psychological distress were used in the analyses of the ALSWH and Whitehall II study, respectively. A further disadvantage of using three cohort studies was the considerable time spent on data management and data cleaning. Since this project did not aim to compare findings of similar analyses across studies, it was decided that the advantages of using three datasets outweighed its disadvantages. However, I acknowledge that findings of this thesis might be specific to the measures and populations used in each of the analyses.

7.3.2 The need for an accurate definition of depression

A major challenge of this project was that depression was not well-defined. As highlighted in the background chapter and evident in the systematic review, there is a wide range of assessment tools for clinical depression and depressive symptoms (see sections 2.1.2 Assessment of depression and 3.4.2.3 Exposure assessment). Due to the large number of potential combinations of diagnostic criteria and the variety of information collected on different depressive symptoms, people who share related symptoms but with different aetiological and biological mechanisms might be grouped together under the term depression.

The issue of defining depression created challenges in my PhD. First, the absence of a gold standard to measure depression or depressive symptoms was evident when I was trying to identify suitable cohort studies to investigate the relationship between depression and CVD. The identified cohort studies assessed depression or depressive symptoms using a wide variety of measurement tools. Furthermore, some of the identified studies changed their measurement of depression during follow-up. Under the assumption that all assessment tools measure the same construct of depression and that all tools identify the same group of participants as depressed, the heterogeneity of measurement tools used would not be problematic. However, as discussed previously (see section 2.1.2.2 Rating scales), it has been highlighted by researchers that different depressive symptom rating scales should not be used interchangeably due to the lack of content overlap. Ideally, I would have liked to use the same measure of depression or depressive symptoms in all of the projects. However, given the variation in measures used across the identified cohort studies, it was not possible to focus on just one depression or depressive symptom measure. As mentioned previously, the results of the projects might therefore be specific to each measure used. It remains to be determined how well the results of this thesis are generalisable to all patients with "depression". As a next step, it will be of interest to determine a preferred measurement tool to assess depression in research. Considering the large input of human and financial resources to conduct interviews to establish the presence of clinical depression, it is unlikely to be feasible to most

cohort studies. Instead, it is very likely that cohort studies will continue to use depressive symptom rating scales or linkage to routinely collected health records. In these cohort studies the use of existing validated instruments should be prioritised over creating new measurement tools. Furthermore, a justification for why a certain tool was used should be presented.

Second, the issue of defining depression had an impact on the interpretation of my results. By relying on depressive symptoms and psychological distress sum scores in the ALSWH and Whitehall II projects, I implicitly made two assumptions (Fried, 2015). First, I assumed that depression is a distinct disease category. Second, I assumed that all symptoms of depression are equally important. However, these assumptions might not be valid. As highlighted by Fried (2017b), certain symptoms of depression might be more important with regard to risk of CVD than others. Thus, my analytic approach simplified a potentially much more complex situation in which, for example, somatic symptoms of depression might be more important than affective symptoms with regard to risk of CVD. Whilst I might have lost potentially important information by not looking at individual items of rating scales, the use of individual items has remained challenging in practice since rating scales were neither developed nor validated to measure individual depressive symptoms (Fried et al, 2016). In future studies, it will be essential to acknowledge the assumptions that are made when relying on sum scores of depressive symptoms or psychological distress.

7.3.3 The challenge of establishing causality between depression and CVD

7.3.3.1 Methods of assessing causality

One conventional way of assessing whether an epidemiologic association between an exposure and outcome of interest is causal is to assess whether is meets the Bradford-Hill criteria of causality (Hill, 1965). The nine proposed criteria are: Strength, consistency, specificity, temporality, biological gradient (dose-response effect), plausibility, coherence, experiment, and analogy (Hill, 1965). Applying this concept to the observed association between clinical depression, depressive symptoms and

risk of CVD, six of the criteria are met. The findings of this PhD support a strong and consistent association, temporality, and potentially support a biological gradient. Furthermore, previous research has reported plausible mechanisms for the observed association and the findings are not in conflict with the general knowledge of the history and biology of both depression and CVD (coherence). Three criteria might not be met. First, so far randomised controlled trials have not been able to show that a change in depressive symptoms results in a change of risk of CVD. Second, the specificity criterion might not be met. Hill (1965) suggested that associations are more likely to be causal when one exposure only causes one disease. In contrast, there are associations between clinical depression or depressive symptoms and a number of different health outcomes, such as all-cause mortality, CVD, and possibly cancer (lack of specificity). However, researcher have argued that the specificity criterion might be more relevant for infectious diseases or chemical exposures than for more complex exposures, such as mental illnesses (Fedak et al, 2015). Third, the analogy criteria specifies that the observed association should be similar to other known causal associations. To the best of my knowledge, no definite causal relationship has been established between severe mental illnesses and cardiovascular diseases yet.

Another more direct way of assessing whether an observed association between an exposure and outcome of interest might be causal is to infer the cause-effect relationship from data. In most cases, data from randomised controlled trials are used to infer such relationship between an intervention and outcome of interest. Some researchers have argued that observational data could be used to infer cause-effect relationships using a counterfactual approach to causation (Hernán & Taubman, 2019). In these situations, observational data are used to mimic randomised controlled trials that might not be feasible, for example due to high financial resources required, or might not be ethical. However, these causal inference methods provide valid estimates of the causal effect under strong assumptions. One such assumption is the consistency assumption. Of note, this consistency assumption (described in more detail below) is not equivalent to the consistency assumption of the Bradford-Hill criteria of causality.

7.3.3.2 The consistency assumption

Covert heterogeneity of depression is particularly problematic if one is interested in causal inference since the term depression might "encompass a multitude of meanings" (Rehkopf et al, 2016, p2). The reason behind this is that exposures with multiple subcomponents and heterogeneous underlying experiences, often referred to as compound treatments, have been shown to violate the consistency assumption (Hernán & VanderWeele, 2011; Rehkopf et al, 2016). The consistency assumption is an important assumption for the application of causal inference methods using the counterfactual approach to causation. It requires that there are no two versions of treatments for which the potential outcome would be different had everyone been exposed to version A or version B (Rehkopf et al, 2016). For example, we might be interested in estimating the causal effect of changing the depressive symptom sum score on risk of CVD. Two potential treatments of depression might be the use of an antidepressant and the use of counselling. Assuming both interventions are equally efficient, they will result in the same change of the depressive symptom sum score. However, both interventions might have direct effects on the outcome of interest that are not mediated through depression. Therefore, the effect of the interventions on CVD might be different even though both interventions resulted in the same change of the depressive symptom sum score. Since the potential outcome would be different had everyone been exposed to the antidepressant or to counselling sessions, the exposure "changes in depressive symptom sum scores" violates the consistency assumption. Acknowledging that all treatments might be somewhat vague, Hernán & Taubman (2008) highlighted that there is a question of degree. It was argued that treatments should be specific enough so that variation between different versions of the treatment is irrelevant to the potential outcome (treatment-variation irrelevance) (Hernán & VanderWeele, 2011). In contrast, researchers should avoid estimating the average causal effect of treatments with multiple versions that differ in their effect on the potential outcome (treatment-variation relevance) (Hernán & VanderWeele, 2011).

7.3.3.3 Causal inference in research on psychological disorders

There is an ongoing discussion about whether or not causal inference methods should be used in research on psychological disorders. Due to its multiple subcomponents and considerable heterogeneity, a violation of the consistency assumption is plausible and likely for depression. Some researchers argue that causal inference methods should not be used in the absence of well-defined interventions (Hernán & Taubman, 2008; Hernán & VanderWeele, 2011). Hernán & VanderWeele (2011) further encourage researchers to imagine a hypothetical RCT, even if it might not be ethical or feasible, to ensure relevant causal questions that fulfil the consistency assumption. Hernán & Taubman (2008) acknowledged that an investigation of the effect of compound exposures may generate interesting hypotheses but they additionally highlight that it remains unclear to what extent estimates can be translated into policies. Other researchers have more nuanced opinions on the application of causal inference methods for compound treatments (Glymour & Kubzansky, 2017; Rehkopf et al, 2016). Rehkopf et al (2016) and Glymour & Kubzansky (2017) stated that the importance of the consistency assumption might be higher once a research area matures whereas it might be less relevant in the early stages of research questions. Glymour & Kubzansky (2017) further highlighted the difficulty of imagining hypothetical interventions in psychosocial epidemiology and emphasised that progress towards a causal understanding might be possible even in the presence of a violation of the consistency assumption. To provide a complete picture, it should be noted that some authors disagree with the concept of consistency as an assumption that needs to be evaluated. Whilst the evaluation of the consistency assumption is an integral part of causal inference based on the counterfactual approach (Hernán & VanderWeele, 2011), other researchers argue that consistency will hold for individuals whether or not the exact version of an exposure is known (Van der Laan et al, 2005), or that it is a logical consequence when using nonparametric structural equation models (Pearl, 2010).

There are additional methodological considerations that need to be taken into account when causal inference methods are used to estimate the effect of depression on risk

of CVD. Since longitudinal studies assess the exposure and covariates repeatedly over time, exposure-covariate feedback (often referred to as treatment-confounder feedback) needs to be considered in the analytic approach (Hernán & Robins, 2019). Treatment-confounder feedback refers to a situation in which covariates affect the exposure and prior exposure status affects the covariate (Figure 7.1). In the presence of treatment-confounder feedback simply adjusting for time-varying covariates leads to biased results since time-varying covariates might act as confounders and mediators at the same time (Glymour et al, 2019). One analytical approach that accounts for potential treatment-confounder feedback is the use of marginal structural models in which the analysis is adjusted for time-invariant covariates whereas inverse probability weighting is used for time-variant confounders. As discussed in section 3.5.2.3 Confounding and mediation, this approach has been used in two published studies on the effect of depression on the risk of CVD (Gilsanz et al, 2017; Gilsanz et al, 2015). Whilst this approach can deal with treatment-confounder feedback, it does not overcome the likely violation of the consistency assumption. One example of exposure-covariate feedback is the relationship between psychotropic medication, depression severity and risk of CVD. This example is further described in section 7.3.4.1 The need to disentangle the role of depression and psychotropic medications.

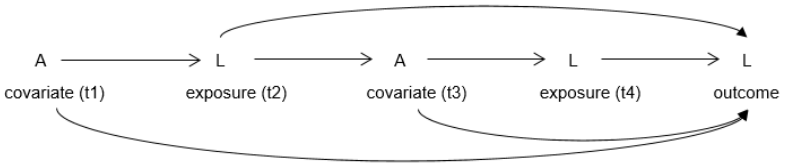


Figure 7.1: Directed acyclic graph of treatment-confounder feedback, adapted from Glymour et al (2019)

NB: Presence of arrow; causal effect is assumed or unwilling to assume that causal effect does not exist; direction of arrow: direction of causal effect

7.3.4 Alternative explanations of the observed association outside the scope of this project

7.3.4.1 The need to disentangle the role of depression and psychotropic medications

Whilst a number of shortcomings of the existing evidence base could be addressed, some alternative explanations of the observed association between depression and CVD remain to be investigated. For example, as discussed in the background section of this thesis, iatrogenic effects of psychotropic medications might partly explain the observed association between depression and CVD. Whilst investigating this alternative explanation was outside the scope of this project, it might result in valuable insights into the relationship between depression and CVD.

Antidepressant and antipsychotic drug use have been shown to be associated with increased risks of MCVE in systematic reviews and meta-analyses. Antidepressant use, defined either as use of any antidepressant, SSRI or tricyclic antidepressant (TCA) alone has been shown to be associated with increased risk of total stroke (Pan et al, 2011b; Shin et al, 2014; Trajkova et al, 2019), ischaemic stroke (Shin et al, 2014), intracranial (Shin et al, 2014) and intracerebral haemorrhage but not subarachnoid stroke (Hackam & Mrkobrada, 2012; Shin et al, 2014). Furthermore, SSRI but not TCA were associated with increased risk of CBVD (Biffi et al, 2017), and TCA but not SSRI were associated with increased risk of acute heart disease (Biffi et al, 2017) and MI (Undela et al, 2015). Antipsychotic drug use has been shown to be associated with increased risks of stroke in two existing meta-analyses (Hsu et al, 2017; Zivkovic et al, 2019) whereas results on the risk of MI among individuals with antipsychotic drug use are conflicting (Brauer et al, 2011; Huang et al, 2017; Yu et al, 2016; Zivkovic et al, 2019). However, most authors emphasised that a cautious interpretation of results is warranted due to high between-study heterogeneity (Biffi et al, 2017; Brauer et al, 2011; Huang et al, 2017; Shin et al, 2014; Zivkovic et al, 2019).

Two of the meta-analyses on the association between antidepressant use and subsequent stroke aimed to estimate the effect of antidepressant use on risk of stroke, independent of depression (Shin et al, 2014; Trajkova et al, 2019). Since an increased

risk remained after restricting their analyses to studies that accounted for potential confounding by depression, the researchers argued for an independent effect of antidepressants. Furthermore, Trajkova et al (2019) highlighted that there was an effect of antidepressant use on risk of stroke in a case-crossover study (Wu et al, 2011) which by design controls for time-invariant confounding. However, it should be noted that the effect of depression is unlikely to be time-invariant and not just depression but also severity of depression needs to be taken into account.

It has been challenging to disentangle the effects of depression and psychotropic medication use since medication use likely acts as an identifier of disease severity at the same time. Therefore, the increased risk among individuals with medication use might reflect increased risk of more severe depression rather than an independent effect of the medication itself. Furthermore, due to treatment-confounder feedback between depression severity and medication use, disease severity is potentially both a confounder and mediator in the relationship between medication use and risk of CVD (see Figure 7.1 for directed acyclic graph). Glymour et al (2019) highlighted the importance of time-updated measures due to the feedback between depression severity and antidepressant use. Potential reasons for the lack of studies incorporating disease severity as potential confounder and mediating factor is the complexity of the analysis as well as the lack of datasets with detailed information on depression severity, medications and outcomes. One notable exception is a recently published study by Glymour et al (2019). Using data from 2,302 adult participants of the Changes In Thoughts cohort study, the authors estimated the effect of ever having initiated antidepressant medication on risk of first-ever stroke over 8.4 years of follow-up. In contrast to existing studies, Glymour et al (2019) adjusted for time-constant confounders and applied inverse probability weights to take account of time-varying covariates, such as disease severity. As highlighted previously, inverse probability weighting is a method to condition on covariates without blocking potential mediating pathways (see section 3.4.2.5 Conditioning on covariates for detailed discussion). Glymour et al (2019) reported that individuals who used antidepressants were at increased risk of stroke. However, this was specific to

patients using TCA but not to patients using SSRI or other antidepressants. A disadvantage of this analysis was that the authors did not account for individuals stopping medication once they initiated it. The authors argued that they aimed to avoid bias due to individuals stopping medications because of side effects or health concerns. Nonetheless, the risk of stroke might differ between individuals who stop and continue taking medications. Therefore, it would be of interest to incorporate this aspect in future research.

7.3.4.2 The need to investigate genetic pleiotropy in depression and CVD

The observed association between depression and CVD might be explained through a genetic region that affects both depression and CVD. This concept is known as genetic pleiotropy (Amare et al, 2016). Genetic pleiotropy has been shown for depression and coronary artery disease (Hagenaars et al, 2019; Wray et al, 2018) and depression and stroke (Hagenaars et al, 2019; Wassertheil-Smoller et al, 2018) in some studies whereas other studies have not found evidence for genetic pleiotropy for depression and CHD (Khandaker et al, 2019) or depression and stroke (Anttila et al, 2018). Although it was outside the scope of this project to investigate the potential common genetic aetiology of depression and CVD, an investigation of this alternative explanation for the observed associations might provide valuable insights into the relationship between depression and CVD. More research into potential shared genetic or biological mechanisms for depression and CVD is needed. However, the success of future studies might depend on a clear and specific definition of depression.

Despite the modest heritability of depression, it has proven difficult to identify independent genetic loci for depression that replicate in other studies. Heritability is defined as the "proportion of total variance in a population for a particular measurement, taken at a particular time or age, that is attributable to variation in additive genetic or total genetic values" (Visscher et al, 2008, p255). By definition, the measure of heritability depends on the study population. The heritability of depression has been estimated to be ~37% in twin studies (Sullivan et al, 2012). The

genome-wide single nucleotide polymorphism (SNP) based heritability of unrelated subjects was estimated to be 8.7% indicating that about one quarter of the heritability of depression is explained by common genetic variants (Wray et al, 2018). Furthermore, some studies suggested that the heritability of depression might differ between different depressive symptoms with somatic symptoms, such as fatigue, having a stronger heritable basis than non-somatic symptoms, such as suicidal ideation (Jang et al, 2004; Thorp et al, 2019). Thorp et al (2019) highlighted that 108 independent genetic loci for schizophrenia had been identified whereas none had been identified for depression by 2014. There are several potential explanations for this null finding. First, the heritability estimates of depression are much lower than those of other severe mental illnesses suggesting that environmental factors might be much more important than genetic risk factors. Second, depression is much more prevalent than other severe mental illnesses (Thorp et al, 2019). Therefore, bigger sample sizes might be needed to identify genetic loci for depression. In keeping with that theory, more recent studies using much larger sample sizes were more successful in identifying genetic loci. For example, Wray et al (2018) identified 44 genetic loci, 30 of which were new findings and 12 of which have been reported in prior studies. Another potential explanation for the lack of findings might be the considerable heterogeneity of depression. Whilst studies relying on bigger sample sizes assume that sample size can overcome clinical heterogeneity (Wray et al, 2018), other researchers have argued that there should be more emphasis on the genetic contribution of unique symptoms (Jang et al, 2004; Thorp et al, 2019).

7.4 Implications for research

The systematic review and meta-analysis convincingly shows that there is an association between a single measure of depression or depressive symptoms and risk of CVD. Furthermore, the results of the UK Biobank analysis suggest that residual confounding alone does not explain the observed association. Considering the results of the systematic review and meta-analysis and the results of the UK Biobank project, there is no need for additional studies on the association between a single measure of depression and subsequent CVD. Instead, future studies should now focus on using

repeat assessments of depression, on investigating the inter-relationship of variables throughout follow-up, and on understanding potential mechanisms involved.

Further research is needed that takes into account the inter-relationship of variables throughout follow-up. For example, a marginal structural model estimating the effect of changes of depressive symptoms on risk of CVD might provide valuable insights into the relationship between depression and CVD. In the ALSWH analysis I showed that there were differences in profiles of key cardiovascular risk factors across subgroups of women with different depressive symptom trajectories at baseline. Furthermore, there were differences in these cardiovascular risk factors over time. Whilst differences between groups at baseline might confound any relationship between depression and risk of CVD, changes in these characteristics during follow-up might act as confounding and/ or mediating factors. Through the use of inverse probability weighting, marginal structural models account for time-varying confounding without blocking potential causal pathways. As explained in more detail previously (see section 7.3.3 The challenge of establishing causality between depression and CVD), the application of causal inference methods in the context of depression is controversial and the analyses are very complex and time-consuming. Furthermore, the likely violation of the consistency assumption should be acknowledged and discussed in future research. A potential approach might be to break down the exposure into more specific subcomponents when trying to evaluate causal effects of depression on risk of CVD. It has been highlighted that researchers should aim to identify specific subcomponents of depression for which, regardless of which component of depression is changed, the same effect would occur (treatment-variation irrelevance) (Glymour & Kubzansky, 2017; Rehkopf et al, 2016). Potential subcomponents might be individual items of depressive symptom rating scales. However, it remains to be determined whether these subcomponents fulfil the consistency assumption.

There is a need for research to now focus on understanding potential mechanisms involved. To the best of my knowledge, no mediation analysis on the mediating effect

of biological or lifestyle factors in the context of depression and CVD has been performed that accounted for treatment-confounder feedback and the likely violation of the consistency assumption. I agree with authors of a previous study that the relationship between depression or depressive symptoms, BMI and subsequent CVD deserves more attention (Jackson & Mishra, 2013). In contrast to findings of the study by Jackson & Mishra (2013) that was based on the mid-aged ALSWH cohort, the majority of women of the young cohort were not overweight or obese at baseline. Interestingly, the mean BMI was similar among women with stable low, stable moderate and fluctuating depressive symptoms at the age of 20 years but the patterns of change of BMI differed from the age of 20 to 40 years. Since the time from the age of 20 to 40 years seems to be important in terms of changes of BMI over time, future studies on the potential mediating effect of BMI should be based on participants that are not older than 20 years at baseline. Since depressive symptoms and BMI are time-varying, there might be feedback between change of BMI and change of depressive symptoms over time. It will be important to acknowledge this feedback in future research and choose an analytic approach, such as inverse probability weighting, that can deal with time-varying confounding without blocking potential mediating pathways.

An additional relationship that deserves more attention is the relationship between depression, physical comorbidities and risk of CVD. It remains to be determined whether physical comorbidities act as common sources and/ or mediators in the relationship between depression and subsequent CVD. In contrast to existing publications, the UK Biobank analysis suggests that individuals with depression but no hypertension are not at increased risk of CVD, relative to those with neither depression nor hypertension, but the risk of individuals with depression and hypertension exceeds the sum of the risks of depression and hypertension alone. The mechanisms explaining this pattern will depend on whether depression or hypertension occurred first (see section 7.2 Contribution of this project to our understanding of the relationship between depression and subsequent CVD). Therefore, it will be of interest to investigate how depression and trajectories of

physical illnesses accrual over time relate to risk of CVD in future studies. Cohort studies with repeat assessments or access to routinely collected health records will be essential in determining what conditions develop first and what conditions follow. In time, one potential cohort study to investigate this relationship further is the young cohort of the ALSWH study as women of this cohort age.

The analysis of data from the Whitehall II study provided an interesting and innovative new perspective on the relationship between psychological distress and cardiovascular events. The next obvious step would be to make use of study phases 9 and 11 for which outcomes have not yet been validated. This would offer advantages in terms of statistical power by increasing the number of events and in terms of follow-up by providing 2.5 years additional years of follow-up in the analysis on fatal CHD or non-fatal MI and five years of additional follow-up in the analysis on fatal or non-fatal stroke. Second, it would be interesting to repeat the analysis using different outcomes such as diagnosis of any CVD or any CHD. Again, the use of a broader outcome definition would increase the number of events in the analysis and thereby increase statistical power to detect differences between trajectories of psychological distress among those with and without events. Third, it remains to be established if results are similar if a measure of depressive symptoms is used instead of a measure of psychological distress. Unfortunately, the current dataset did not allow for a replication of findings using a measure of depressive symptoms due to lack of validated cardiovascular events in study phases 9 and 11. Once outcomes are validated for these waves and additional assessments of depressive symptoms are collected, it will be of interest to compare the findings based on psychological distress with findings based on depressive symptoms. Fourth, given that this was the first study that compared psychological distress trajectories prior to diagnosis with cardiovascular events to psychological distress trajectories among individuals free from cardiovascular events over the same time period, it would be of interest to replicate the analysis using a different dataset. Ideally, this dataset would have equally extensive information on the exposure before diagnosis of cardiovascular events and a long-follow-up period that covers the pre-clinical phase of CVD. A

cohort study that offers similar advantages is the ALSHW. Since the number of women that had a stroke and myocardial infarction was very small due to the young age of the young cohort of the ALSWH at their last follow-up assessment, depressive symptom trajectories prior to diagnosis with hypertension could be compared to depressive symptom trajectories among individuals free from hypertension over the same time period. Once women of the young cohort of the ALSWH were followed up for a longer time and more cardiovascular events have occurred, the analysis could be repeated using a diagnosis of stroke and/ or MI as outcome. Alternatively, data from the mid-aged ALSWH cohort could be used to replicate the Whitehall II analysis.

Lastly, further research is needed to investigate other alternative explanations of the observed associations that were beyond the scope of this project. As described in more detail previously (see sections 7.3.4.1 The need to disentangle the role of depression and psychotropic medications and 7.3.4.2 The need to investigate genetic pleiotropy in depression and CVD), two potential alternative explanations that warrant further research are iatrogenic effects of psychotropic medications and genetic pleiotropy in depression and CVD.

7.5 Implications for practice

This project furthered our understanding of the relationship between depression and CVD. Yet, it remains to be established to what extent clinical depression and depressive symptoms are indeed causally related with increased risk of CVD. Whilst further research is clearly needed, this section proposes possible implications for practice based on the findings of this PhD.

Targeted screening or preventive programmes for CVD and its risk factors may be needed, if successful treatment strategies are available to those that are identified as being at high risk of CVD. Since findings of this thesis suggest that individuals with depression and physical comorbidities and/ or individuals with depression from low socioeconomic backgrounds are at particularly high risk of CVD, screening strategies should be targeted at these individuals. There are existing health check programmes

for CVD and its risk factors delivered in primary care in England and Scotland (National Health Service, 2016b; NHS Health Scotland, no date). The NHS health check programme in England is a universal programme targeted at all adults aged 40 – 74 years without pre-existing diabetes or CVD. In contrast to the English health check programme, the Scottish programme Keep Well uses a targeted approach to identify adults aged 40 – 64 years at particularly high risk of CVD. Furthermore, there was a particular emphasis on reducing health inequalities and targeting individuals living in areas of Scotland with high deprivation. Both health check programmes were evaluated four years after implementation. The overall coverage of the English programme was low (Chang et al, 2015; Robson et al, 2016), there was no significant difference in coverage by deprivation (Chang et al, 2015), a small effect on early identification of CVD and risk factor management (Chang et al, 2019) but this effect did not differ by deprivation (Chang et al, 2019). Thus, the English health check programme was not successful in reducing health inequalities. The Keep Well programme successfully targeted individuals in the most deprived areas (NHS Health Scotland, 2014). However, the effects of the Keep Well programme on incident hospitalisations, prescription rates, and CHD and stroke mortality were small (Geue et al, 2016). To increase the effect of future preventive programmes at a population level, these programmes should incorporate recommendations of NHS Health Scotland (2014). For example, both the English and Scottish programmes offered brief advice on changes of health behaviours or prescription of relevant medication for individuals at high risk of CVD. Instead, future programme should offer continuous support to individuals and, more importantly, should incorporate changes to the living environment which takes away the responsibility from individuals.

To decrease health inequalities between individuals with depression from low and high socioeconomic backgrounds with regard to risk of CVD, an alignment of general practice funding with the social gradient in clinical need is urgently needed (McLean et al, 2015; Mercer & Watt, 2007). In the UK, the majority of health care services are provided free of charge to all UK residents. Whilst this ensures access to care, McLean et al (2015) highlighted that it does not ensure that the provision and funding of

general practice is aligned with the high clinical need in deprived areas. In keeping with that Mercer & Watt (2007) showed that the increased need in areas with high area-based deprivation was associated with "poorer access to care, less time spent with the doctor, higher GP stress, and lower patient enablement in encounters for psychosocial problems" (Mercer & Watt, 2007, p507). To overcome this problem, the funding allocation formula for general practice may need to be altered. McLean et al (2015) suggested that the allocation of funding for general practice should consider that individuals living in areas of high deprivation are at higher risk of multimorbidity from early adulthood onwards. The authors emphasised that general practices in more deprived areas will be disadvantaged if the allocation of funding is based on the age of the practice population.

The results of the ALSWH analysis suggest that more effective primary and secondary prevention strategies are needed for risky alcohol intake and weight gain whereas smoking prevention was effective among women with stable low, stable moderate and fluctuating depressive symptoms. Since the proportion of current smokers among individuals with stable low, stable moderate and fluctuating depressive symptoms decreased over time, the results of the ALSWH analysis suggest that primary and/ or secondary prevention programmes are effective with regards to smoking. In contrast, the proportion of individuals with risky/ high risk alcohol intake increased over time. Furthermore, women with stable moderate and fluctuating depressive symptoms gained weight more rapidly than women with stable low depressive symptoms. Thus, these results suggest that there is a need to develop more effective strategies to prevent risky alcohol intake and weight gain. Based on results of the ASLWH analysis it seems particularly important to monitor and prevent weight gain among individuals with stable moderate and fluctuating depressive symptoms in their 20s. Potential explanations for weight gain among women with stable moderate and fluctuating depressive symptoms are an unhealthy diet or low levels of physical activity. However, a further explanation may be that these women are more likely to use antidepressants which in turn may promote weight gain.

Irrespective of the potential independent effect of depression on CVD onset and progression, efforts towards better recognition and treatment of depression are warranted. Despite its high prevalence and considerable contribution to the global burden of diseases, it has been shown that the majority of people with MDD do not receive adequate treatment (Thornicroft et al, 2017). The reasons behind this are multifactorial. First, individuals with depression might not recognise their need for treatment. Using data from the WHO World Mental Health Surveys from 21 countries, Thornicroft et al (2017) reported that only 56.7% of all individuals who met the DSM-IV criteria of MDD in the past 12 months recognised their need for treatment. Therefore, there is a need to encourage individuals with depressive symptoms to seek medical help and to increase awareness that depression is a treatable condition. However, there is also a need to improve the supply of timely and effective services since it has been shown that individuals who recognise their need for treatment might not receive treatment. Thornicroft et al (2017) reported that 71.1% of all participants who recognised their need for treatment made at least one visit to a service provider but only 41.0% received treatment meeting minimal standards. Since it has been shown that recovery rates of individuals with mental health problems are associated with shorter waiting times between referral and treatment, higher numbers of treatment sessions, and better ability of service providers to identify the problem they are treating, service level factors should additionally be taken into account to achieve better treatment of individuals with depression (Clark et al, 2018).

Chapter 8: Conclusion

Although an association between depression and CVD onset and progression has been observed in numerous primary studies and existing reviews, there was an ongoing debate whether or not the evidence base was sufficient to acknowledge depression as an independent risk factor of CVD. This was due to potential non-causal explanations of the observed association and methodological shortcomings of existing studies.

This PhD furthered our understanding of the relationship between depression and CVD by exploring the merit of two potential non-causal explanation of the observed association. The results of the UK Biobank analysis suggest that it is unlikely that residual confounding due to unmeasured covariates alone explains the observed associations between different measure of depression and MCVE. The results of the Whitehall II analysis suggest that it is unlikely that reverse causation or overlap of symptoms between subclinical CVD and somatic symptoms of depression explains the observed association. In addition, this project highlights the importance of using repeat assessments of depressive symptoms and identifies individuals with depression who are at particularly high risk of CVD.

Considering the substantial contribution of both depression and CVD to the global burden of disease, it should remain a public health priority to further investigate the relationship between depression and CVD. Whilst this PhD furthered our understanding of the relationship between depression and subsequent CVD, more research is clearly needed to establish depression as an independent risk factor of CVD. Instead of investigating the relationship between a single measure of depression and risk of CVD, research should now focus on using repeat assessments of depression, on investigating the inter-relationship of variables throughout follow-up, and on understanding potential mechanisms involved.

Bibliography

Air, T., Tully, P. J., Sweeney, S. & Beltrame, J. (2016) Epidemiology of Cardiovascular Disease and Depression, *Cardiovascular Diseases and Depression*. Springer, 5-21.

Albus, C., Herrmann-Lingen, C., Jensen, K., Hackbusch, M., Münch, N., Kuncewicz, C., Grilli, M., Schwaab, B., Rauch, B. & German Society of Cardiovascular Prevention & Rehabilitation (2019) Additional effects of psychological interventions on subjective and objective outcomes compared with exercise-based cardiac rehabilitation alone in patients with cardiovascular disease: a systematic review and meta-analysis. *European Journal of Preventive Cardiology*, 1-15.

Alexopoulos, G. S., Meyers, B. S., Young, R. C., Campbell, S., Silbersweig, D. & Charlson, M. (1997) 'Vascular depression' hypothesis. *Archives Of General Psychiatry*, 54(10), 915-922.

Allen, N., Sudlow, C., Downey, P., Peakman, T., Danesh, J., Elliott, P., Gallacher, J., Green, J., Matthews, P., Pell, J., Sprosen, T. & Collins, R. (2012) UK Biobank: Current status and what it means for epidemiology. *Health Policy and Technology*, 1(3), 123-126.

Amare, A. T., Schubert, K. O. & Baune, B. T. (2016) Genetic Overlap Between Depression and Cardiometabolic Disorders, *Cardiovascular Diseases and Depression*. Springer, 235-255.

American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders : DSM-5*, Fifth edition. Arlington, VA : American Psychiatric Association.

Ames, M. E. & Leadbeater, B. J. (2018) Depressive symptom trajectories and physical health: Persistence of problems from adolescence to young adulthood. *Journal of Affective Disorders*, 240, 121-129.

Andersson, T., Alfredsson, L., Kallberg, H., Zdravkovic, S. & Ahlbom, A. (2005) Calculating measures of biological interaction. *European Journal of Epidemiology*, 20(7), 575-579.

Andresen, E. M., Malmgren, J. A., Carter, W. B. & Patrick, D. L. (1994) Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). *American Journal of Preventive Medicine*, 10(2), 77-84.

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., Escott-Price, V., Falcone, G. J., Gormley, P., Malik, R., Patsopoulos, N. A., Ripke, S., Wei, Z., Yu, D., Lee, P. H., Turley, P., Grenier-Boley, B., Chouraki, V., Kamatani, Y., Berr, C., Letenneur, L., Hannequin, D., Amouyel, P., Boland, A., Deleuze, J.-F., Duron, E., Vardarajan, B. N., Reitz, C., Goate, A. M., Huentelman, M. J., Kamboh, M. I., Larson, E. B., Rogaeva, E., St George-Hyslop, P., Hakonarson, H., Kukull, W. A., Farrer, L. A., Barnes, L. L., Beach, T. G., Demirci, F. Y., Head, E., Hulette, C. M., Jicha, G. A., Kauwe, J. S. K., Kaye, J. A., Leverenz, J. B., Levey, A. I., Lieberman, A. P., Pankratz, V. S., Poon, W. W., Quinn, J. F., Saykin, A. J., Schneider, L. S., Smith, A. G., Sonnen, J. A., Stern, R. A., Van Deerlin, V. M., Van Eldik, L. J., Harold, D., Russo, G., Rubinsztein, D. C., Bayer, A., Tsolaki, M., Proitsi, P., Fox, N. C., Hampel, H., Owen, M. J., Mead, S., Passmore, P., Morgan, K., Nöthen, M. M., Schott, J. M., Rossor, M., Lupton, M. K., Hoffmann, P., Kornhuber, J., Lawlor, B., McQuillin, A.,

Al-Chalabi, A., Bis, J. C., Ruiz, A., Boada, M., Seshadri, S., Beiser, A., Rice, K., van der Lee, S. J., De Jager, P. L., Geschwind, D. H., Riemenschneider, M., Riedel-Heller, S., Rotter, J. I., Ransmayr, G., Hyman, B. T., Cruchaga, C., Alegret, M., Winsvold, B., Palta, P., Farh, K.-H., Cuenca-Leon, E., Furlotte, N., et al (2018) Analysis of shared heritability in common disorders of the brain. *Science*, 360(6395), eaap8757.

Arbelaez, J. J., Ariyo, A. A., Crum, R. M., Fried, L. P. & Ford, D. E. (2007) Depressive symptoms, inflammation, and ischemic stroke in older adults: a prospective analysis in the cardiovascular health study. *Journal of the American Geriatrics Society*, 55(11), 1825-1830.

Ariyo, A. A., Haan, M., Tangen, C. M., Rutledge, J. C., Cushman, M., Dobs, A. & Furberg, C. D. (2000) Depressive symptoms and risks of coronary heart disease and mortality in elderly Americans. *Circulation*, 102(15), 1773-1779.

Avendano, M., Kawachi, I., Van Lenthe, F., Boshuizen, H. C., Mackenbach, J. P., Van Den Bos, G. A. M., Fay, M. E. & Berkman, L. F. (2006) Socioeconomic status and stroke incidence in the US elderly: the role of risk factors in the EPESE study. *Stroke*, 37(6), 1368-1373.

Ayerbe, L., Forgnone, I., Addo, J., Siguero, A., Gelati, S. & Ayis, S. (2018) Hypertension risk and clinical care in patients with bipolar disorder or schizophrenia; a systematic review and meta-analysis. *Journal of Affective Disorders*, 225, 665-670.

Barefoot, J. C. & Schroll, M. (1996) Symptoms of depression, acute myocardial infarction, and total mortality in a community sample. *Circulation*, 93(11), 1976-1980.

Barlinn, K., Kepplinger, J., Puetz, V., Illigens, B. M., Bodechtel, U. & Siepmann, T. (2014) Exploring the risk-factor association between depression and incident stroke: A systematic review and meta-analysis. *Neuropsychiatric Disease and Treatment*, 11(1), 1-14.

Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S. & Guthrie, B. (2012) Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet*, 380(9836), 37-43.

Batty, G. D., McIntosh, A. M., Russ, T. C., Deary, I. J. & Gale, C. R. (2016) Psychological distress, neuroticism, and cause-specific mortality: early prospective evidence from UK Biobank. *Journal of Epidemiology and Community Health*, 70(11), 1136-1139.

Baxter, A. J., Charlson, F. J., Somerville, A. J. & Whiteford, H. A. (2011) Mental disorders as risk factors: assessing the evidence for the Global Burden of Disease Study. *BMC Medicine*, 9, 134-134.

Beatriz Currier, M. & B Nemeroff, C. (2010) Inflammation and mood disorders: proinflammatory cytokines and the pathogenesis of depression. *Anti-Inflammatory & Anti-Allergy Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Inflammatory and Anti-Allergy Agents)*, 9(3), 212-220.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J. & Erbaugh, J. (1961) An Inventory for Measuring Depression. *Archives of General Psychiatry*, 4(6), 561-571.

Benjamin, E. J., Muntner, P., Alonso, A., Bittencourt, M. S., Callaway, C. W., Carson, A. P., Chamberlain, A. M., Chang, A. R., Cheng, S., Das, S. R., Delling, F. N.,

Djousse, L., Elkind, M. S. V., Ferguson, J. F., Fornage, M., Jordan, L. C., Khan, S. S., Kissela, B. M., Knutson, K. L., Kwan, T. W., Lackland, D. T., Lewis, T. T., Lichtman, J. H., Longenecker, C. T., Loop, M. S., Lutsey, P. L., Martin, S. S., Matsushita, K., Moran, A. E., Mussolino, M. E., O'Flaherty, M., Pandey, A., Perak, A. M., Rosamond, W. D., Roth, G. A., Sampson, U. K. A., Satou, G. M., Schroeder, E. B., Shah, S. H., Spartano, N. L., Stokes, A., Tirschwell, D. L., Tsao, C. W., Turakhia, M. P., VanWagner, L. B., Wilkins, J. T., Wong, S. S. & Virani, S. S. (2019) Heart Disease and Stroke Statistics - 2019 Update: A Report From the American Heart Association. *Circulation*, 139(10), e56-e528.

Biffi, A., Scotti, L. & Corrao, G. (2017) Use of antidepressants and the risk of cardiovascular and cerebrovascular disease: a meta-analysis of observational studies. *European Journal of Clinical Pharmacology*, 73(4), 487-497.

Bjelland, I., Dahl, A. A., Haug, T. T. & Neckelmann, D. (2002) The validity of the Hospital Anxiety and Depression Scale: An updated literature review. *Journal of Psychosomatic Research*, 52(2), 69-77.

Bobo, W. V., Yawn, B. P., St. Sauver, J. L., Grossardt, B. R., Boyd, C. M. & Rocca, W. A. (2016) Prevalence of Combined Somatic and Mental Health Multimorbidity: Patterns by Age, Sex, and Race/Ethnicity. *The Journals of Gerontology: Series A*, 71(11), 1483-1491.

Bos, M. J., Lindén, T., Koudstaal, P. J., Hofman, A., Skoog, I., Breteler, M. M. B. & Tiemeier, H. (2008) Depressive symptoms and risk of stroke: the Rotterdam Study. *Journal of Neurology, Neurosurgery & Psychiatry*, 79(9), 997-1001.

Boucquemont, J., Loubere, L., Metzger, M., Combe, C., Stengel, B. & Leffondre, K. (2017) Identifying subgroups of renal function trajectories. *Nephrology Dialysis Transplantation*, 32(suppl_2), ii185-ii193.

Brauer, R., Douglas, I. & Smeeth, L. (2011) The association between antipsychotic agents and the risk of myocardial infarction: a systematic review. *British Journal of Clinical Pharmacology*, 72(6), 871-878.

Brown, J. M., Stewart, J. C., Stump, T. E. & Callahan, C. M. (2011) Risk of coronary heart disease events over 15 years among older adults with depressive symptoms. *The American Journal of Geriatric Psychiatry*, 19(8), 721-729.

Brown, W. J., Dobson, A. J., Bryson, L. & Byles, J. E. (1999) Women's Health Australia: On the Progress of the Main Cohort Studies. *Journal of Women's Health & Gender-Based Medicine*, 8(5), 681-688.

Brown, W. J. & Pavey, T. (2016) Physical Activity in Mid-Age and Older Women: Lessons from the Australian Longitudinal Study on Women's Health. *Kinesiology Review*, 5(1), 87-97.

Brunner, E. J., Shipley, M. J., Britton, A. R., Stansfeld, S. A., Heuschmann, P. U., Rudd, A. G., Wolfe, C. D., Singh-Manoux, A. & Kivimaki, M. (2014) Depressive disorder, coronary heart disease, and stroke: dose-response and reverse causation effects in the Whitehall II cohort study. *European Journal of Preventive Cardiology*, 21(3), 340-346.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A. & Kaemmer, B. (1989) *MMPI-2: Minnesota Multiphasic Personality Inventory-2. Manual for administration and scoring.* Minneapolis, MN: University of Minnesota Press.

Capistrant, B. D., Gilsanz, P., Moon, J. R., Kosheleva, A., Patton, K. K. & Glymour, M. M. (2013) Does the association between depressive symptoms and cardiovascular mortality risk vary by race? Evidence from the Health and Retirement Study. *Ethnicity & Disease*, 23(2), 155-160.

Carpenter, J. S., Andrykowski, M. A., Wilson, J., Hall, L. A., Rayens, M. K., Sachs, B. & Cunningham, L. L. (1998) Psychometrics for two short forms of the Center for Epidemiologic Studies-Depression Scale. *Issues in Mental Health Nursing*, 19(5), 481-494.

Carroll, A. J. M., Auer, R., Colangelo, L. A., Carnethon, M. R., Jacobs, D. R., Jr., Stewart, J. C., Widome, R., Carr, J. J., Liu, K. & Hitsman, B. (2017) Association of the Interaction Between Smoking and Depressive Symptom Clusters With Coronary Artery Calcification: The CARDIA Study. *Journal of Dual Diagnosis*, 13(1), 43-51.

Cassidy, W. L., Flanagan, N. B., Spellman, M. & Cohen, M. E. (1957) Clinical observations in manic-depressive disease: A quantitative study of one hundred manic-depressive patients and fifty medically sick controls. *Journal of the American Medical Association*, 164(14), 1535-1546.

Chang, K. C.-M., Soljak, M., Lee, J. T., Woringer, M., Johnston, D., Khunti, K., Majeed, A. & Millett, C. (2015) Coverage of a national cardiovascular risk assessment and management programme (NHS Health Check): Retrospective database study. *Preventive Medicine*, 78, 1-8.

Chang, K. C.-M., Vamos, E. P., Palladino, R., Majeed, A., Lee, J. T. & Millett, C. (2019) Impact of the NHS Health Check on inequalities in cardiovascular disease risk: a difference-in-differences matching analysis. *Journal of Epidemiology and Community Health*, 73(1), 11-18.

Charlson, F. J., Baxter, A. J., Whiteford, H. A. & Stapelberg, N. J. C. (2011) Should Global Burden of Disease Estimates Include Depression as a Risk Factor for Coronary Heart Disease? *BMC Medicine*, 9(1), 47.

Charlson, F. J., Moran, A. E., Freedman, G., Norman, R. E., Stapelberg, N. J. C., Baxter, A. J., Vos, T. & Whiteford, H. A. (2013) The contribution of major depression to the global burden of ischemic heart disease: a comparative risk assessment. *BMC Medicine*, 11(1), 250.

Chi, M. J., Yu, E., Liu, W. W., Lee, M. C. & Chung, M. H. (2014) The bidirectional relationship between myocardial infarction and depressive disorders: a follow-up study. *International Journal of Cardiology*, 177(3), 854-859.

Clarivate Analytics (2016) *EndNote X8.0.1*. Philadelphia, USA. Available online: https://endnote.com/ [Accessed 12/12/2019].

Clark, D. M., Canvin, L., Green, J., Layard, R., Pilling, S. & Janecka, M. (2018) Transparency about the outcomes of mental health services (IAPT approach): an analysis of public data. *The Lancet*, 391(10121), 679-686.

Cohen, B. E., Edmondson, D. & Kronish, I. M. (2015) State of the art review: depression, stress, anxiety, and cardiovascular disease. *American Journal of Hypertension*, 28(11), 1295-1302.

Colman, I. & Ataullahjan, A. (2010) Life Course Perspectives on the Epidemiology of Depression. *The Canadian Journal of Psychiatry*, 55(10), 622-632.

Connor Gorber, S., Schofield-Hurwitz, S., Hardt, J., Levasseur, G. & Tremblay, M. (2009) The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine & Tobacco Research*, 11(1), 12-24.

Correll, C. U., Solmi, M., Veronese, N., Bortolato, B., Rosson, S., Santonastaso, P., Thapa-Chhetri, N., Fornaro, M., Gallicchio, D., Collantoni, E., Pigato, G., Favaro, A., Monaco, F., Kohler, C., Vancampfort, D., Ward, P. B., Gaughran, F., Carvalho, A. F. & Stubbs, B. (2017) Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry*, 16(2), 163-180.

Costello, D. M., Swendsen, J., Rose, J. S. & Dierker, L. C. (2008) Risk and protective factors associated with trajectories of depressed mood from adolescence to early adulthood. *J Consult Clin Psychol*, 76(2), 173-83.

Critical Appraisal Skills Programme (2019) *CASP Cohort Study Checklist [online].* Available online: https://casp-uk.net/wp-content/uploads/2018/01/CASP-Cohort-Study-Checklist_2018.pdf [Accessed 15/06/2017].

Crump, C., Sundquist, K., Winkleby, M. A. & Sundquist, J. (2013) Comorbidities and Mortality in Bipolar Disorder: A Swedish National Cohort Study. *JAMA Psychiatry*, 70(9), 931-939.

Cummings, D. M., Kirian, K., Howard, G., Howard, V., Ya, Y., Muntner, P., Kissela, B., Redmond, N., Judd, S. E., Safford, M. M. & Yuan, Y. (2016) Consequences of Comorbidity of Elevated Stress and/or Depressive Symptoms and Incident Cardiovascular Outcomes in Diabetes: Results From the REasons for Geographic And Racial Differences in Stroke (REGARDS) Study. *Diabetes Care*, 39(1), 101-109.

Curran, P. J. & Muthén, B. O. (1999) The Application of Latent Curve Analysis to Testing Developmental Theories in Intervention Research. *American Journal of Community Psychology*, 27(4), 567-595.

Dal-Ré, R., Ioannidis, J. P., Bracken, M. B., Buffler, P. A., Chan, A.-W., Franco, E. L., La Vecchia, C. & Weiderpass, E. (2014) Making Prospective Registration of Observational Research a Reality. *Science Translational Medicine*, 6(224), 224cm1.

Daskalopoulou, M., George, J., Walters, K., Osborn, D. P., Batty, G. D., Stogiannis, D., Rapsomaniki, E., Pujades-Rodriguez, M., Denaxas, S., Udumyan, R., Kivimaki, M. & Hemingway, H. (2016) Depression as a Risk Factor for the Initial Presentation of Twelve Cardiac, Cerebrovascular, and Peripheral Arterial Diseases: Data Linkage Study of 1.9 Million Women and Men. *PLOS ONE*, 11(4), e0153838.

de Geus, E. J. C. (2006) Genetic pleiotropy in depression and coronary artery disease. *Psychosomatic Medicine*, 68(2), 185-186.

Dickersin, K. & Min, Y. I. (1993) Publication bias: the problem that won't go away. *Annals of the New York Academy of Sciences*, 703, 135-146.

DiMatteo, M., Lepper, H. S. & Croghan, T. W. (2000) Depression is a risk factor for noncompliance with medical treatment: Meta-analysis of the effects of anxiety and depression on patient adherence. *Archives of Internal Medicine*, 160(14), 2101-2107.

Dong, J. Y., Zhang, Y. H., Tong, J. & Qin, L. Q. (2012) Depression and risk of stroke: A meta-analysis of prospective studies. *Stroke*, 43(1), 32-37.

Doyle, F., McGee, H., Conroy, R., Conradi, H. J., Meijer, A., Steeds, R., Sato, H., Stewart, D. E., Parakh, K. & Carney, R. (2015) Systematic review and individual patient data meta-analysis of sex differences in depression and prognosis in persons with myocardial infarction: a MINDMAPS study. *Psychosomatic Medicine*, 77(4), 419-428.

Everson-Rose, S. A., Roetker, N. S., Lutsey, P. L., Kershaw, K. N., Longstreth, W. T., Jr., Sacco, R. L., Diez Roux, A. V. & Alonso, A. (2014) Chronic stress, depressive symptoms, anger, hostility, and risk of stroke and transient ischemic attack in the multi-ethnic study of atherosclerosis. *Stroke*, 45(8), 2318-2323.

Everson, S. A., Roberts, R. E., Goldberg, D. E. & Kaplan, G. A. (1998) Depressive symptoms and increased risk of stroke mortality over a 29-year period. *Archives of Internal Medicine*, 158(10), 1133-1138.

Fan, Z., Wu, Y., Shen, J., Ji, T. & Zhan, R. (2013) Schizophrenia and the risk of cardiovascular diseases: A meta-analysis of thirteen cohort studies. *Journal of Psychiatric Research*, 47(11), 1549-1556.

Fedak, K. M., Bernal, A., Capshaw, Z. A. & Gross, S. (2015) Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. *Emerging Themes in Epidemiology*, 12(1), 14.

Feighner, J. P., Robins, E., Guze, S. B., Woodruff, R. A., Winokur, G. & Munoz, R. (1972) Diagnostic Criteria for Use in Psychiatric Research. *Archives of General Psychiatry*, 26(1), 57-63.

Ferrari, A. J., Norman, R. E., Freedman, G., Baxter, A. J., Pirkis, J. E., Harris, M. G., Page, A., Carnahan, E., Degenhardt, L., Vos, T. & Whiteford, H. A. (2014) The burden attributable to mental and substance use disorders as risk factors for suicide: findings from the global burden of disease study 2010. *PLOS ONE*, 9(4), e91936.

Fewell, Z., Hernán, M. A., Wolfe, F., Tilling, K., Choi, H. & Sterne, J. A. C. (2004) Controlling for Time-dependent Confounding using Marginal Structural Models. *The Stata Journal*, 4(4), 402-420.

Fiedorowicz, J. G. (2014) Depression and cardiovascular disease: an update on how course of illness may influence risk. *Current Psychiatry Reports*, 16(10), 492.

First, M., Williams, J., Karg, R. & Spitzer, R. (2015) *Structured Clinical Interview for DSM-5, Research version (SCID-5 for DSM-5, Research Version; SCID-5-RV)*. Arlington, VA: American Psychiatric Association.

Fried, E. I. (2015) Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. *Frontiers in Psychology* 6, 309.

Fried, E. I. (2017a) The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191-197.

Fried, E. I. (2017b) Moving forward: how depression heterogeneity hinders progress in treatment and research. *Expert Review of Neurotherapeutics*, 17(5), 423-425.

Fried, E. I. & Nesse, R. M. (2015a) Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, 172, 96-102.

Fried, E. I. & Nesse, R. M. (2015b) Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Medicine*, 13(1), 72.

Fried, E. I., Nesse, R. M., Zivin, K., Guille, C. & Sen, S. (2014) Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychological Medicine*, 44(10), 2067-2076.

Fried, E. I., Van Borkulo, C. D., Epskamp, S., Schoevers, R. A., Tuerlinckx, F. & Borsboom, D. (2016) Measuring Depression Over Time . . . or not? Lack of Unidimensionality and Longitudinal Measurement Invariance in Four Common Rating Scales of Depression. *Psychological Assessment*, 28(11), 1354-1367.

Furukawa, T., Anraku, K., Hiroe, T., Takahashi, K., Kitamura, T., Hirai, T., Takahashi, K. & Iida, M. (1997) Screening for depression among first-visit psychiatric patients: comparison of different scoring methods for the Center for Epidemiologic Studies Depression Scale using receiver operating characteristic analyses. *Psychiatry and Clinical Neurosciences*, 51(2), 71-78.

Gafarov, V. V., Gromova, E. A., Gagulin, I. V., Panov, D. O. & Gafarova, A. V. (2017) Gender peculiarities of the risk of cardiovascular diseases in a population with symptoms of depression in Siberia (the WHO MONICA-psychosocial program). *Terapevticheskii Arkhiv*, 89(9), 60-67.

Gan, Y., Gong, Y., Tong, X., Sun, H., Cong, Y., Dong, X., Wang, Y., Xu, X., Yin, X., Deng, J., Li, L., Cao, S. & Lu, Z. (2014) Depression and the risk of coronary heart disease: a meta-analysis of prospective cohort studies. *BMC Psychiatry*, 14(1), 371.

Gatov, E., Rosella, L., Chiu, M. & Kurdyak, P. A. (2017) Trends in standardized mortality among individuals with schizophrenia, 1993-2012: a population-based, repeated cross-sectional study. *Canadian Medical Association Journal*, 189(37), e1177-e1187.

Gaysina, D., Hotopf, M., Richards, M., Colman, I., Kuh, D. & Hardy, R. (2011) Symptoms of depression and anxiety, and change in body mass index from adolescence to adulthood: results from a British birth cohort. *Psychological Medicine*, 41(1), 175-184.

GBD 2015 Disease and Injury Incidence and Prevalence Collaborators (2016) Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 388(10053), 1545-1602.

GBD 2015 Mortality and Causes of Death Collaborators (2016) Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, 388(10053), 1459-1544.

Geue, C., Lewsey, J. D., MacKay, D. F., Antony, G., Fischbacher, C. M., Muirie, J. & McCartney, G. (2016) Scottish Keep Well health check programme: an interrupted time series analysis. *Journal of Epidemiology and Community Health*, 70(9), 924-929.

Gilsanz, P., Kubzansky, L. D., Tchetgen, E. J. T., Wang, Q. Y., Kawachi, I., Patton, K. K., Fitzpatrick, A. L., Kop, W. J., Longstreth, W. T. & Glymour, M. M. (2017) Changes in Depressive Symptoms and Subsequent Risk of Stroke in the Cardiovascular Health Study. *Stroke*, 48(1), 43-48.

Gilsanz, P., Walter, S., Tchetgen, E. J. T., Patton, K. K., Moon, J. R., Capistrant, B. D., Marden, J. R., Kubzansky, L. D., Kawachi, I. & Glymour, M. M. (2015) Changes in

Depressive Symptoms and Incidence of First Stroke Among Middle-Aged and Older US Adults. *Journal of the American Heart Association*, 4(5), e001923.

Global Burden of Disease Collaborative Network (2018) *Global Burden of Disease Study 2017 (GBD 2017) Results.* Seattle, WA: Institute for Health Metrics and Evaluation (IHME). Available online: http://ghdx.healthdata.org/gbd-results-tool [Accessed 01/05/2018].

Glymour, M., Maselko, J., Gilman, S., Patton, K. & Avendano, M. (2010) Depressive symptoms predict incident stroke independently of memory impairments. *Neurology*, 75(23), 2063-2070.

Glymour, M., Yen, J., Kosheleva, A., Moon, J., Capistrant, B. & Patton, K. (2012) Elevated depressive symptoms and incident stroke in Hispanic, African-American, and White older Americans. *Journal of Behavioral Medicine*, 35(2), 211-220.

Glymour, M. M., Gibbons, L. E., Gilsanz, P., Gross, A. L., Mez, J., Brewster, P. W., Marden, J., Zahodne, L. B., Nho, K., Hamilton, J., Li, G., Larson, E. B. & Crane, P. K. (2019) Initiation of antidepressant medication and risk of incident stroke: using the Adult Changes in Thought cohort to address time-varying confounding. *Annals of Epidemiology*, 35, 42-47.

Glymour, M. M. & Kubzansky, L. D. (2017) Causal inference in psychosocial epidemiology: Key concepts and methods, in Kivimäki, M., Batty, D. G., Kawachi, I. & Steptoe, A. (eds), *The Routledge International Handbook of Psychosocial Epidemiology*. New York: Routledge.

Goldberg, D. P. (1972) *The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness*. Oxford, England: Oxford U. Press.

Goldston, K. & Baillie, A. J. (2008) Depression and coronary heart disease: a review of the epidemiological evidence, explanatory mechanisms and management approaches. *Clinical Psychology Review*, 28(2), 288-306.

Gonzalez, J. S., Peyrot, M., McCarl, L. A., Collins, E. M., Serpa, L., Mimiaga, M. J. & Safren, S. A. (2008) Depression and Diabetes Treatment Nonadherence: A Meta-Analysis. *Diabetes Care*, 31(12), 2398-2403.

Gump, B. B., Matthews, K. A., Eberly, L. E. & Chang, Y.-f. (2005) Depressive symptoms and mortality in men: results from the Multiple Risk Factor Intervention Trial. *Stroke*, 36(1), 98-102.

Gustad, L. T., Laugsand, L. E., Janszky, I., Dalen, H. & Bjerkeset, O. (2014) Symptoms of anxiety and depression and risk of acute myocardial infarction: the HUNT 2 study. *European Heart Journal*, 35(21), 1394-403.

Hackam, D. G. & Mrkobrada, M. (2012) Selective serotonin reuptake inhibitors and brain hemorrhage: a meta-analysis. *Neurology*, 79(18), 1862-1865.

Hagenaars, S., Coleman, J., Choi, S., Gaspar, H., Adams, M., Howard, D., Hodgson, K., Traylor, M., Air, T., Andlauer, T., Arolt, V., Baune, B., Binder, E., Blackwood, D., Boomsma, D., Campbell, A., Cearns, M., Czamara, D., Dannlowski, U., Domschke, K., de Geus, E., Hamilton, S., Hayward, C., Hickie, I., Hottenga, J., Ising, M., Jones, I., Jones, L., Kutalik, Z., Lucae, S., Martin, N., Milaneschi, Y., Mueller-Myhsok, B., Owen, M., Padmanabhan, S., Penninx, B., Pistis, G., Porteous,

D., Preisig, M., Ripke, S., Shyn, S., Sullivan, P., Whitfield, J., Wray, N., McIntosh, A., Deary, I., Breen, G. & Lewis, C. (2019) Genetic comorbidity between major depression and cardio-metabolic disease, stratified by age at onset of major depression. *bioRxiv*, 645077.

Hatano, S. (1976) Experience from a multicentre stroke register: a preliminary report. *Bulletin of the World Health Organization*, 54(5), 541-553.

Hathaway, S. R. & McKinley, J. C. (1951) *Minnesota Multiphasic Personality Inventory; Manual, revised*. San Antonio, TX, US: Psychological Corporation.

Hawkins, M. A., Callahan, C. M., Stump, T. E. & Stewart, J. C. (2014) Depressive symptom clusters as predictors of incident coronary artery disease events: A 15-year prospective study of older adults. *Psychosomatic Medicine*, 76(1), 38-43.

Head, J., Stansfeld, S. A., Ebmeier, K. P., Geddes, J. R., Allan, C. L., Lewis, G. & Kivimaki, M. (2013) Use of self-administered instruments to assess psychiatric disorders in older people: validity of the General Health Questionnaire, the Center for Epidemiologic Studies Depression Scale and the self-completion version of the revised Clinical Interview Schedule. *Psychological Medicine*, 43(12), 2649-2656.

Henderson, K. M., Clark, C. J., Lewis, T. T., Aggarwal, N. T., Beck, T., Guo, H., Lunos, S., Brearley, A., Mendes de Leon, C. F., Evans, D. A. & Everson-Rose, S. A. (2013) Psychosocial distress and stroke risk in older adults. *Stroke*, 44(2), 367-72.

Hernán, M. A., Brumback, B. & Robins, J. M. (2000) Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*, 11(5), 561-70.

Hernán, M. A. & Robins, J. M. (2019) Treatment-confounder feedback, *Causal Inference*, forthcoming edition. Boca Raton: Chapman & Hall/CRC.

Hernán, M. A. & Taubman, S. L. (2008) Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *International Journal of Obesity*, 32(S3), 8-14.

Hernán, M. A. & Taubman, S. L. (2019) Observational studies, *Causal Inference*, forthcoming edition. Boca Raton: Chapman & Hall/CRC.

Hernán, M. A. & VanderWeele, T. J. (2011) Compound treatments and transportability of causal inference. *Epidemiology*, 22(3), 368-377.

Higgins, J. P. & Green, S. (2011) *Cochrane handbook for systematic reviews of interventions Version 5.1.0 [updated March 2011]*. Oregon Health & Science University, Portland, Oregon: The Cochrance Collaboration.

Hill, A. B. (1965) The environment and disease: association or causation? *Journal of the Royal Society of Medicine*, 108(1), 32-37.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D. G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S. E., Dixon-Woods, M., McCulloch, P., Wyatt, J. C., Chan, A.-W. & Michie, S. (2014) Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ : British Medical Journal*, 348, g1687.

Holden, L., Ware, R. S. & Lee, C. (2016) Trajectories of mental health over 16 years amongst young adult women: The Australian Longitudinal Study on Women's Health. *Developmental Psychology*, 52(1), 164-75.

Hsu, W. T., Esmaily-Fard, A., Lai, C. C., Zala, D., Lee, S. H., Chang, S. S. & Lee, C. C. (2017) Antipsychotics and the Risk of Cerebrovascular Accident: A Systematic Review and Meta-Analysis of Observational Studies. *Journal of the American Medical Directors Association*, 18(8), 692-699.

Huang, C.-J., Hsieh, M.-H., Hou, W.-H., Liu, J.-C., Jeng, C. & Tsai, P.-S. (2013) Depression, antidepressants, and the risk of coronary heart disease: a population-based cohort study. *International Journal of Cardiology*, 168(5), 4711-4716.

Huang, K. L., Fang, C. J., Hsu, C. C., Wu, S. I., Juang, J. J. & Stewart, R. (2017) Myocardial infarction risk and antipsychotics use revisited: a meta-analysis of 10 observational studies. *Journal of Psychopharmacology*, 31(12), 1544-1555.

IPAQ Group (2010) *Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) - Short Form.* London: IPAQ Group. Available online: https://sites.google.com/site/theipaq/scoring-protocol [Accessed 23/10/2017].

Jackson, C. (2007) The General Health Questionnaire. *Occupational Medicine*, 57(1), 79-79.

Jackson, C. A. & Mishra, G. D. (2013) Depression and Risk of Stroke in Midaged Women A Prospective Longitudinal Study. *Stroke*, 44(6), 1555-1560.

Jackson, C. A., Pathirana, T. & Gardiner, P. A. (2016) Depression, anxiety and risk of hypertension in mid-aged women: a prospective longitudinal study. *Journal of Hypertension*, 34(10), 1959-1966.

Jackson, C. A., Sudlow, C. L. & Mishra, G. D. (2018) Psychological Distress and Risk of Myocardial Infarction and Stroke in the 45 and Up Study: A Prospective Cohort Study. *Circulation-Cardiovascular Quality and Outcomes*, 11(9), e004500.

Jang, K. L., Livesley, W. J., Taylor, S., Stein, M. B. & Moon, E. C. (2004) Heritability of individual depressive symptoms. *Journal of Affective Disorders*, 80(2-3), 125-133.

Janszky, I., Ahnve, S., Lundberg, I. & Hemmingsson, T. (2010) Early-onset depression, anxiety, and risk of subsequent coronary heart disease: 37-year follow-up of 49,321 young Swedish men. *Journal of the American College of Cardiology*, 56(1), 31-37.

Jørgensen, M., Mainz, J., Carinci, F., Thomsen, R. W. & Johnsen, S. P. (2017) Quality and Predictors of Diabetes Care Among Patients With Schizophrenia: A Danish Nationwide Study. *Psychiatric Services*, 69(2), 179-185.

Joyce, C. (2015) *Exploring the relationship between anxiety and depression in the onset and recovery of coronary heart disease : A mixed methods study*. PhD Dissertation. Western Sydney University. Available online: https://researchdirect.westernsydney.edu.au/islandora/object/uws%3A40667 [Accessed 02/02/2019].

Kapfhammer, H.-P. (2006) Somatic symptoms in depression. *Dialogues in Clinical Neuroscience*, 8(2), 227-239.

Kapur, S., Phillips, A. G. & Insel, T. R. (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Molecular Psychiatry*, 17(12), 1174-1179.

Kemp, A. H., Quintana, D. S., Gray, M. A., Felmingham, K. L., Brown, K. & Gatt, J. M. (2010) Impact of depression and antidepressant treatment on heart rate variability: a review and meta-analysis. *Biological Psychiatry*, 67(11), 1067-1074.

Khandaker, G. M., Zuber, V., Rees, J. M. B., Carvalho, L., Mason, A. M., Foley, C. N., Gkatzionis, A., Jones, P. B. & Burgess, S. (2019) Shared mechanisms between coronary heart disease and depression: findings from a large UK general population-based cohort. *Molecular Psychiatry*.

Kim, E. S., Park, N. & Peterson, C. (2011) Dispositional optimism protects older adults from stroke: The health and retirement study. *Stroke*, 42(10), 2855-2859.

Kim, E. S., Sun, J. K., Park, N. & Peterson, C. (2013) Purpose in life and reduced incidence of stroke in older adults: 'The Health and Retirement Study'. *Journal of Psychosomatic Research*, 74(5), 427-432.

Köhler, S., Verhey, F., Weyerer, S., Wiese, B., Heser, K., Wagner, M., Pentzek, M., Fuchs, A., Köhler, M., Bachmann, C., Riedel Heller, S. G., Luppa, M., Eifflaender-Gorfer, S., Werle, J., Bickel, H., Mösch, E., König, H.-H., Brettschneider, C., Scherer, M. & Maier, W. (2013) Depression, non-fatal stroke and all-cause mortality in old age: A prospective cohort study of primary care patients. *Journal of Affective Disorders*, 150(1), 63-69.

Kubzansky, L. D., Cole, S. R., Kawachi, I., Vokonas, P. & Sparrow, D. (2006) Shared and unique contributions of anger, anxiety, and depression to coronary heart disease: A prospective study in the normative aging study. *Annals of Behavioral Medicine*, 31(1), 21-29.

Kubzansky, L. D., Davidson, K. W. & Rozanski, A. (2005) The clinical impact of negative psychological states: expanding the spectrum of risk for coronary artery disease. *Psychosomatic Medicine*, 67 Suppl 1, S10-14.

Kuller, L. H., Arnold, A. M., Psaty, B. M., Robbins, J. A., O'Leary, D. H., Tracy, R. P., Burke, G. L., Manolio, T. A. & Chaves, P. H. (2006) 10-year follow-up of subclinical cardiovascular disease and risk of coronary heart disease in the Cardiovascular Health Study. *Archives of Internal Medicine*, 166(1), 71-8.

Kurdyak, P., Vigod, S., Duchen, R., Jacob, B., Stukel, T. & Kiran, T. (2017) Diabetes quality of care and outcomes: Comparison of individuals with and without schizophrenia. *General Hospital Psychiatry*, 46, 7-13.

Kutlubaev, M. A. & Hackett, M. L. (2014) Part II: Predictors of Depression after Stroke and Impact of Depression on Stroke Outcome: An Updated Systematic Review of Observational Studies. *International Journal of Stroke*, 9(8), 1026-1036.

Langvik, E. & Hjemdal, O. (2015) Symptoms of depression and anxiety before and after myocardial infarction: The HUNT 2 and HUNT 3 study. *Psychology, Health & Medicine*, 20(5), 560-569.

Larson, S. L., Owens, P. L., Ford, D. & Eaton, W. (2001) Depressive disorder, dysthymia, and risk of stroke: thirteen-year follow-up from the Baltimore epidemiologic catchment area study. *Stroke*, 32(9), 1979-1983.

Lawrence, D., Hancock, K. J. & Kisely, S. (2013) The gap in life expectancy from preventable physical illness in psychiatric patients in Western Australia: retrospective analysis of population based registers. *BMJ*, 346, f2539.

Lazzarino, A. I., Hamer, M., Stamatakis, E. & Steptoe, A. (2013a) The combined association of psychological distress and socioeconomic status with all-cause mortality: a national cohort study. *JAMA Internal Medicine*, 173(1), 22-27.

Lazzarino, A. I., Hamer, M., Stamatakis, E. & Steptoe, A. (2013b) Low socioeconomic status and psychological distress as synergistic predictors of mortality from stroke and coronary heart disease. *Psychosomatic Medicine*, 75(3), 311-316.

Lee, C., Dobson, A. J., Brown, W. J., Bryson, L., Byles, J., Warner-Smith, P. & Young, A. F. (2005) Cohort Profile: The Australian Longitudinal Study on Women's Health. *International Journal of Epidemiology*, 34(5), 987-991.

Leigh, L., Byles, J. E., Chojenta, C. & Pachana, N. A. (2016) Late life changes in mental health: a longitudinal study of 9683 women. *Aging & Mental Health*, 20(10), 1044-1054.

Lennon, H., Kelly, S., Sperrin, M., Buchan, I., Cross, A. J., Leitzmann, M., Cook, M. B. & Renehan, A. G. (2018) Framework to construct and interpret latent class trajectory modelling. *BMJ Open*, 8(7), e020683.

Leung, Y. W., Flora, D. B., Gravely, S., Irvine, J., Carney, R. M. & Grace, S. L. (2012) The impact of premorbid and postmorbid depression onset on mortality and cardiac morbidity among patients with coronary heart disease: meta-analysis. *Psychosomatic Medicine*, 74(8), 786-801.

Lewis, G., Pelosi, A. J., Araya, R. & Dunn, G. (1992) Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychological Medicine*, 22(2), 465-486.

Li, M., Fan, Y.-L., Tang, Z.-Y. & Cheng, X.-S. (2014) Schizophrenia and risk of stroke: A meta-analysis of cohort studies. *International Journal of Cardiology*, 173(3), 588-590.

Li, M., Zhang, X. W., Hou, W. S. & Tang, Z. Y. (2015) Impact of depression on incident stroke: a meta-analysis. *International Journal of Cardiology*, 180, 103-110.

Liang, J., Xu, X., Quinones, A. R., Bennett, J. M. & Ye, W. (2011) Multiple trajectories of depressive symptoms in middle and late life: racial/ethnic variations. *Psychol Aging*, 26(4), 761-77.

Licht, C. M., de Geus, E. J. & Penninx, B. W. (2013) Dysregulation of the autonomic nervous system predicts the development of the metabolic syndrome. *The Journal of Clinical Endocrinology and Metabolism*, 98(6), 2484-2493.

Lichtman, J. H., Froelicher, E. S., Blumenthal, J. A., Carney, R. M., Doering, L. V., Frasure-Smith, N., Freedland, K. E., Jaffe, A. S., Leifheit-Limson, E. C., Sheps, D. S., Vaccarino, V. & Wulsin, L. (2014) Depression as a Risk Factor for Poor Prognosis Among Patients With Acute Coronary Syndrome: Systematic Review and Recommendations: A Scientific Statement From the American Heart Association. *Circulation*, 129(12), 1350-1369.

Lin, Y.-N., Lin, C.-L., Chang, Y.-J., Peng, C.-L., Sung, F.-C., Chang, K.-C. & Kao, C.-H. (2014) Increased subsequent risk of acute coronary syndrome for patients with depressive disorder: A nationwide population-based retrospective cohort study. *Psychiatry and Clinical Neurosciences*, 68(4), 263-269.

Linden, W., Phillips, M. J. & Leclerc, J. (2007) Psychological treatment of cardiac patients: a meta-analysis. *European Heart Journal*, 28(24), 2972-2984.

Liu, N. H., Daumit, G. L., Dua, T., Aquila, R., Charlson, F., Cuijpers, P., Druss, B., Dudek, K., Freeman, M., Fujii, C., Gaebel, W., Hegerl, U., Levav, I., Munk Laursen, T., Ma, H., Maj, M., Elena Medina-Mora, M., Nordentoft, M., Prabhakaran, D.,

Pratt, K., Prince, M., Rangaswamy, T., Shiers, D., Susser, E., Thornicroft, G., Wahlbeck, K., Fekadu Wassie, A., Whiteford, H. & Saxena, S. (2017) Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas. *World Psychiatry*, 16(1), 30-40.

Lo, Y., Mendell, N. R. & Rubin, D. B. (2001) Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767-778.

Loxton, D., Tooth, L., Harris, M. L., Forder, P. M., Dobson, A., Powers, J., Brown, W., Byles, J. & Mishra, G. (2017) Cohort Profile: The Australian Longitudinal Study on Women's Health (ALSWH) 1989–95 cohort. *International Journal of Epidemiology*, 47(2), 391-392e.

Lustman, P. J., Anderson, R. J., Freedland, K. E., de Groot, M., Carney, R. M. & Clouse, R. E. (2000) Depression and poor glycemic control: a meta-analytic review of the literature. *Diabetes Care*, 23(7), 934-942.

Lux, V. & Kendler, K. S. (2010) Deconstructing major depression: A validation study of the DSM-IV symptomatic criteria. *Psychological Medicine*, 40(10), 1679-1690.

Magaard, J. L., Seeralan, T., Schulz, H. & Brütt, A. L. (2017) Factors associated with help-seeking behaviour among individuals with major depression: A systematic review. *PLOS ONE*, 12(5), e0176730.

Majed, B., Arveiler, D., Bingham, A., Ferrieres, J., Ruidavets, J. B., Montaye, M., Appleton, K., Haas, B., Kee, F., Amouyel, P., Ducimetiere, P. & Empana, J. P. (2012) Depressive symptoms, a time-dependent risk factor for coronary heart disease and stroke in middle-aged men: The PRIME study. *Stroke*, 43(7), 1761-1767.

Marijnissen, R. M., Wouts, L., Schoevers, R. A., Bremmer, M. A., Beekman, A. T., Comijs, H. C. & Oude Voshaar, R. C. (2014) Depression in context of low neuroticism is a risk factor for stroke: a 9-year cohort study. *Neurology*, 83(19), 1692-1698.

Marmot, M. & Brunner, E. (2005) Cohort Profile: The Whitehall II study. *International Journal of Epidemiology*, 34(2), 251-256.

Marmot, M. G., Smith, G. D., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E. & Feeney, A. (1991) Health inequalities among British civil servants: the Whitehall II study. *The Lancet*, 337(8754), 1387-1393.

Martin, D. J., Ul-Haq, Z., Nicholl, B. I., Cullen, B., Evans, J., Gill, J. M. R., Roberts, B., Gallacher, J., Mackay, D., McIntosh, A., Hotopf, M., Craddock, N., Deary, I. J., Pell, J. P. & Smith, D. J. (2016) Cardiometabolic disease and features of depression and bipolar disorder: population-based, cross-sectional study. *The British Journal of Psychiatry*, 208(4), 343-351.

Mathur, R., Perez-Pinar, M., Foguet-Boreu, Q., Ayis, S. & Ayerbe, L. (2016) Risk of incident cardiovascular events amongst individuals with anxiety and depression: A prospective cohort study in the east London primary care database. *Journal of Affective Disorders*, 206, 41-47.

McCormick, N., Lacaille, D., Bhole, V. & Avina-Zubieta, J. A. (2014) Validity of Myocardial Infarction Diagnoses in Administrative Databases: A Systematic Review. *PLOS ONE*, 9(3), e92286.

McLean, G., Gunn, J., Wyke, S., Guthrie, B., Watt, G. C., Blane, D. N. & Mercer, S. W. (2014) The influence of socioeconomic deprivation on multimorbidity at different ages: a cross-sectional study. *British Journal of General Practice*, 64(624), e440-447.

McLean, G., Guthrie, B., Mercer, S. W. & Watt, G. C. (2015) General practice funding underpins the persistence of the inverse care law: cross-sectional study in Scotland. *British Journal of General Practice*, 65(641), e799-805.

Mejía-Lancheros, C., Estruch, R., Martinez-Gonzalez, M. A., Salas-Salvado, J., Castaner, O., Corella, D., Aros, F., Gomez-Gracia, E., Fiol, M., Lapetra, J., Serra-Majem, L., Pinto, X., Ros, E., Diez-Espino, J., Basora, J., Sorli, J. V., Lamuela-Raventos, R. M., Ruiz-Gutierrez, V., Munoz, M. A. & Predimed Study Investigators (2014) Impact of psychosocial factors on cardiovascular morbimortality: a prospective cohort study. *BMC Cardiovascular Disorders*, 14(1), 9.

Melchior, M., Chastang, J. F., Head, J., Goldberg, M., Zins, M., Nabi, H. & Younes, N. (2013) Socioeconomic position predicts long-term depression trajectory: a 13-year follow-up of the GAZEL cohort study. *Mol Psychiatry*, 18(1), 112-21.

Mendis, S., Thygesen, K., Kuulasmaa, K., Giampaoli, S., Mahonen, M., K., N. B., Lisheng, L. & Writing group on behalf of the participating experts of the WHO consultation for revision of WHO definition of myocardial infarction (2010) World Health Organisation definition of myocardial infarction: 2008–09 revision. *International Journal of Epidemiology*, 40(1), 139-146.

Meng, L., Chen, D., Yang, Y., Zheng, Y. & Hui, R. (2012) Depression increases the risk of hypertension incidence: a meta-analysis of prospective cohort studies. *Journal of Hypertension*, 30(5), 842-851.

Mercer, S. W. & Watt, G. C. M. (2007) The inverse care law: clinical primary care encounters in deprived and affluent areas of Scotland. *Annals of Family Medicine*, 5(6), 503-510.

Mittag, O. & Meyer, T. (2012) The association of depressive symptoms and ischemic heart disease in older adults is not moderated by gender, marital status or education. *International Journal of Public Health*, 57(1), 79-85.

Mohebbi, M., Nguyen, V., McNeil, J. J., Woods, R. L., Nelson, M. R., Shah, R. C., Storey, E., Murray, A. M., Reid, C. M., Kirpach, B., Wolfe, R., Lockery, J. E. & Berk, M. (2018) Psychometric properties of a short form of the Center for Epidemiologic Studies Depression (CES-D-10) scale for screening depressive symptoms in healthy community dwelling older adults. *General Hospital Psychiatry*, 51, 118-125.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G. & PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA StatementThe PRISMA Statement. *Annals of Internal Medicine*, 151(4), 264-269.

Moise, N., Khodneva, Y., Richman, J., Shimbo, D., Kronish, I. & Safford, M. M. (2016) Elucidating the Association Between Depressive Symptoms, Coronary Heart Disease, and Stroke in Black and White Adults: The REasons for Geographic And Racial Differences in Stroke (REGARDS) Study. *Journal of the American Heart Association*, 5(8), e003767.

Moran, A. E., Forouzanfar, M. H., Roth, G., Mensah, G. A., Ezzati, M., Flaxman, A., Murray, C. J. L. & Naghavi, M. (2014) The Global Burden of Ischemic Heart Disease

in 1990 and 2010: The Global Burden of Disease 2010 Study. *Circulation*, 129(14), 1493-1501.

Mortimer, J. T., Shanahan, M. J. & Kirkpatrick Johnson, M. (2016) *Handbook of the life course Volume II*. Minnesota, MN, USA: Springer Science & Business Media.

Mumford, E. A., Liu, W., Hair, E. C. & Yu, T.-C. (2013) Concurrent trajectories of BMI and mental health patterns in emerging adulthood. *Social Science & Medicine*, 98, 1-7.

Musliner, K. L., Munk-Olsen, T., Eaton, W. W. & Zandi, P. P. (2016) Heterogeneity in long-term trajectories of depressive symptoms: Patterns, predictors and outcomes. *Journal of Affective Disorders*, 192, 199-211.

Nabi, H., Singh-Manoux, A., Shipley, M., Gimeno, D., Marmot, M. G. & Kivimaki, M. (2008) Do psychological factors affect inflammation and incident coronary heart disease: the Whitehall II Study. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 28(7), 1398-406.

Nagin, D. S. (2005) *Group-Based Modeling of Development*Harvard University Press.

Nagin, D. S. & Odgers, C. L. (2010) Group-based trajectory modeling in clinical research. *Annual Review of Clinical Psychology*, 6, 109-38.

National Health and Medical Research Council (2001) *Australian alcohol guidelines: health risks and benefits.* Canberra, Australia: National Health and Medical Research Council. Available online: https://trove.nla.gov.au/work/9859301?q&versionId=44907256 [Accessed 15/02/2018].

National Health Service (2016a) *Clinical depression.* Available online: https://www.nhs.uk/conditions/clinical-depression/ [Accessed 15/11/2016].

National Health Service (2016b) *NHS Health Check.* Available online: https://www.nhs.uk/conditions/nhs-health-check/?tabname=menu [Accessed 24/09/19].

National Health Service (2018) *Cardiovascular diseases.* Available online: https://www.nhs.uk/conditions/cardiovascular-disease/ [Accessed 19/06/2019].

Navrady, L. B., Wolters, M. K., MacIntyre, D. J., Clarke, T. K., Campbell, A. I., Murray, A. D., Evans, K. L., Seckl, J., Haley, C., Milburn, K., Wardlaw, J. M., Porteous, D. J., Deary, I. J. & McIntosh, A. M. (2018) Cohort Profile: Stratifying Resilience and Depression Longitudinally (STRADL): a questionnaire follow-up of Generation Scotland: Scottish Family Health Study (GS:SFHS). *International Journal of Epidemiology*, 47(1), 13-14.

Nemeroff, C. B. & Goldschmidt-Clermont, P. J. (2012) Heartache and heartbreak-the link between depression and cardiovascular disease. *Nature Reviews Cardiology*, 9(9), 526-539.

Newcomer, J. W. & Hennekens, C. H. (2007) Severe mental illness and risk of cardiovascular disease. *JAMA*, 298(15), 1794-1796.

NHS Health Scotland (2014) *The impact of Keep Well: An evaluation of the Keep Well programme from 2006 to 2012.* Edinburgh: NHS Health Scotland. Available online: http://www.healthscotland.com/uploads/documents/23893-Keep_Well_Impact_Evaluation_report.pdf [Accessed 21/09/2019].

NHS Health Scotland (no date) *Keep Well.* Available online: http://www.healthscotland.com/keep-well.aspx [Accessed 24/09/2019].

Nicholson, A., Kuper, H. & Hemingway, H. (2006) Depression as an aetiologic and prognostic factor in coronary heart disease: a meta-analysis of 6362 events among 146 538 participants in 54 observational studies. *European Heart Journal*, 27(23), 2763-2774.

Niles, A. N. & O'Donovan, A. (2018) Comparing anxiety and depression to obesity and smoking as predictors of major medical illnesses and somatic symptoms. *Health Psychology*, 38(2), 172-181.

Nordentoft, M., Wahlbeck, K., Hallgren, J., Westman, J., Osby, U., Alinaghizadeh, H., Gissler, M. & Laursen, T. M. (2013) Excess mortality, causes of death and life expectancy in 270,770 patients with recent onset of mental disorders in Denmark, Finland and Sweden. *PLOS ONE*, 8(1), e55176.

O'Brien, E. C., Greiner, M. A., Sims, M., Hardy, N. C., Wang, W., Shahar, E., Hernandez, A. F. & Curtis, L. H. (2015) Depressive Symptoms and Risk of Cardiovascular Events in Blacks Findings From the Jackson Heart Study. *Circulation-Cardiovascular Quality and Outcomes*, 8(6), 552-559.

Ohira, T., Iso, H., Satoh, S., Sankai, T., Tanigawa, T., Ogawa, Y., Imano, H., Sato, S., Kitamura, A. & Shimamoto, T. (2001) Prospective study of depressive symptoms and risk of stroke among Japanese. *Stroke*, 32(4), 903-906.

Orme, J. G., Reis, J. & Herz, E. J. (1986) Factorial and Discriminant Validity of the Center for Epidemiological Studies Depression (CES-D) Scale. *Journal of Clinical Psychology*, 42(1), 28-33.

Page, M. J., Shamseer, L. & Tricco, A. C. (2018) Registration of systematic reviews in PROSPERO: 30,000 records and counting. *Systematic Reviews*, 7(1), 32.

Pan, A., Okereke, O. I., Sun, Q., Logroscino, G., Manson, J. E., Willett, W. C., Ascherio, A., Hu, F. B. & Rexrode, K. M. (2011a) Depression and Incident Stroke in Women. *Stroke*, 42(10), 2770-2775.

Pan, A., Sun, Q., Okereke, O. I., Rexrode, K. M. & Hu, F. B. (2011b) Depression and risk of stroke morbidity and mortality: a meta-analysis and systematic review. *JAMA*, 306(11), 1241-1249.

Pearl, J. (2010) On the consistency rule in causal inference: axiom, definition, assumption, or theorem? *Epidemiology*, 21(6), 872-875.

Pedersen, S. S. & Doyle, F. (2019) Effectiveness of psychological intervention as add-on to standard cardiac rehabilitation: Time to adopt new methods or keep doing more of the same? *European Journal of Preventive Cardiology*, 26(10), 1032-1034.

Penninx, B. W. (2016a) Mechanisms Linking Depression to Cardiovascular Disease: What Do Epidemiological Studies Tell Us?, in Baune, B. & Tully, P. (eds), *Cardiovascular Diseases and Depression*. Switzerland: Springer, Cham, 37-52.

Penninx, B. W. J. H. (2016b) Depression and cardiovascular disease: Epidemiological evidence on their linking mechanisms. *Neuroscience and Biobehavioral Reviews*, 74(Part B), 277-286.

Péquignot, R., Dufouil, C., Prugger, C., Peres, K., Artero, S., Tzourio, C. & Empana, J.-P. (2016) High level of depressive symptoms at repeated study visits and risk of

coronary heart disease and stroke over 10 years in older adults: The three-city study. *Journal of the American Geriatrics Society*, 64(1), 118-125.

Péquignot, R., Tzourio, C., Peres, K., Ancellin, M. L., Perier, M. C., Ducimetiere, P. & Empana, J. P. (2013) Depressive symptoms, antidepressants and disability and future coronary heart disease and stroke events in older adults: the Three City Study. *European Journal of Epidemiology*, 28(3), 249-256.

Pogosova, N., Saner, H., Pedersen, S. S., Cupples, M. E., McGee, H., Höfer, S., Doyle, F., Schmid, J.-P. & von Känel, R. (2015) Psychosocial aspects in cardiac rehabilitation: From theory to practice. A position paper from the Cardiac Rehabilitation Section of the European Association of Cardiovascular Prevention and Rehabilitation of the European Society of Cardiology. *European Journal of Preventive Cardiology*, 22(10), 1290-1306.

Pössel, P., Mitchell, A. M., Ronkainen, K., Kaplan, G. A., Kauhanen, J. & Valtonen, M. (2015) Do depressive symptoms predict the incidence of myocardial infarction independent of hopelessness? *Journal of Health Psychology*, 20(1), 60-68.

Pratt, L. A., Ford, D. E., Crum, R. M., Armenian, H. K., Gallo, J. J. & Eaton, W. W. (1996) Depression, psychotropic medication, and risk of myocardial infarction. Prospective data from the Baltimore ECA follow-up. *Circulation*, 94(12), 3123-3129.

Prieto, M. L., Cuéllar-Barboza, A. B., Bobo, W. V., Roger, V. L., Bellivier, F., Leboyer, M., West, C. P. & Frye, M. A. (2014) Risk of myocardial infarction and stroke in bipolar disorder: a systematic review and exploratory meta-analysis. *Acta Psychiatrica Scandinavica*, 130(5), 342-353.

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor Gorber, S. & Tremblay, M. (2008) A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 5(1), 56.

Proust-Lima, C., Dartigues, J. & Jacqmin-Gadda, H. (2011) Misuse of the Linear Mixed Model When Evaluating Risk Factors of Cognitive Decline. *American Journal of Epidemiology*, 174(9), 1077-1088.

Proust-Lima, C., Philipps, V. & Liquet, B. (2017) Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package lcmm. *2017*, 78(2), 1-56.

Rådholm, K., Wiréhn, A. B., Chalmers, J. & Östgren, C. J. (2016) Use of antidiabetic and antidepressant drugs is associated with increased risk of myocardial infarction: a nationwide register study. *Diabetic Medicine*, 33(2), 218-223.

Radloff, L. S. (1977) The CES-D scale a self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385-401.

Ramsay, J. O. (1988) Monotone Regression Splines in Action. *Statist. Sci.*, 3(4), 425-441.

Reavell, A. J., Hopkinson, A. M., Clarkesmith, A. D. & Lane, A. D. (2018) Effectiveness of Cognitive Behavioral Therapy for Depression and Anxiety in Patients With Cardiovascular Disease: A Systematic Review and Meta-Analysis. *Psychosomatic Medicine*, 80(8), 742-753.

Rehkopf, D. H., Glymour, M. M. & Osypuk, T. L. (2016) The Consistency Assumption for Causal Inference in Social Epidemiology: When a Rose is Not a Rose. *Current Epidemiology Reports*, 3(1), 63-71.

Reisinger Walker, E., McGee, R. E. & Druss, B. G. (2015) Mortality in mental disorders and global disease burden implications: A systematic review and meta-analysis. *JAMA Psychiatry*, 72(4), 334-341.

Richards, S. H., Anderson, L., Jenkinson, C. E., Whalley, B., Rees, K., Davies, P., Bennett, P., Liu, Z., West, R., Thompson, D. R. & Taylor, R. S. (2017) Psychological interventions for coronary heart disease. *Cochrane Database of Systematic Reviews*(4), Art. No.: CD002902.

Roberts, T., Miguel Esponda, G., Krupchanka, D., Shidhaye, R., Patel, V. & Rathod, S. (2018) Factors associated with health service utilisation for common mental disorders: a systematic review. *BMC Psychiatry*, 18(1), 262.

Robins, J. M., Hernan, M. A. & Brumback, B. (2000) Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5), 550-60.

Robson, J., Dostal, I., Sheikh, A., Eldridge, S., Madurasinghe, V., Griffiths, C., Coupland, C. & Hippisley-Cox, J. (2016) The NHS Health Check in England: an evaluation of the first 4 years. *BMJ Open*, 6(1), e008840.

Roth, A. G., Huffman, D. M., Moran, E. A., Feigin, A. V., Mensah, J. L. G., Naghavi, J. L. M. & Murray, J. L. C. (2015a) Global and Regional Patterns in Cardiovascular Mortality From 1990 to 2013. *Circulation*, 132(17), 1667-1678.

Roth, G. A., Johnson, C. O., Nguyen, G., Naghavi, M., Feigin, V. L., Murray, C. J. L., Forouzanfar, M. H. & Vos, T. (2015b) Methods for Estimating the Global Burden of Cerebrovascular Diseases. *Neuroepidemiology*, 45(3), 146-151.

Rugulies, R. (2002) Depression as a predictor for coronary heart disease: a review and meta-analysis. *American Journal of Preventive Medicine*, 23(1), 51-61.

Sacco, R. L., Kasner, S. E., Broderick, J. P., Caplan, L. R., Connors, J. J., Culebras, A., Elkind, M. S. V., George, M. G., Hamdan, A. D., Higashida, R. T., Hoh, B. L., Janis, L. S., Kase, C. S., Kleindorfer, D. O., Lee, J.-M., Moseley, M. E., Peterson, E. D., Turan, T. N., Valderrama, A. L. & Vinters, H. V. (2013) An Updated Definition of Stroke for the 21st Century: A Statement for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke*, 44(7), 2064-2089.

Saha, S., Chant, D. & McGrath, J. (2007) A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? *Archives of General Psychiatry*, 64(10), 1123-1131.

Scherrer, J. F., Chrusciel, T., Zeringue, A., Garfield, L. D., Hauptman, P. J., Lustman, P. J., Freedland, K. E., Carney, R. M., Bucholz, K. K., Owen, R. & True, W. R. (2010) Anxiety disorders increase risk for incident myocardial infarction in depressed and nondepressed Veterans Administration patients. *American Heart Journal*, 159(5), 772-779.

Scherrer, J. F., Garfield, L. D., Chrusciel, T., Hauptman, P. J., Carney, R. M., Freedland, K. E., Owen, R., True, W. R. & Lustman, P. J. (2011) Increased risk of myocardial infarction in depressed patients with type 2 diabetes. *Diabetes Care*, 34(8), 1729-1734.

Schnier, C., Sudlow, C. & UK Biobank Cardiac Outcomes Adjudication Group (2017a) *Definitions of acute myocardial infarction (MI) and main MI pathological types for UK Biobank phase 1 outcomes adjudication (Version 1).* Available online:

https://biobank.ctsu.ox.ac.uk/crystal/docs/alg_outcome_mi.pdf [Accessed 15/03/2017].

Schnier, C., Sudlow, C. & UK Biobank Stroke Outcomes Adjudication Group (2017b) *Definitions of stroke and main stroke pathological types for UK Biobank phase 1 outcomes adjudication (Version 1).* Available online: https://biobank.ctsu.ox.ac.uk/crystal/docs/alg_outcome_stroke.pdf [Accessed 15/03/2017].

Schubert, K. O., Clark, S. R., Van, L. K., Collinson, J. L. & Baune, B. T. (2017) Depressive symptom trajectories in late adolescence and early adulthood: A systematic review. *Australian & New Zealand Journal of Psychiatry*, 51(5), 477-499.

Schwarzer, G. (2007) meta: An R package for meta-analysis. *R News*, 7(3), 40-45.

Sesso, H. D., Kawachi, I., Vokonas, P. S. & Sparrow, D. (1998) Depression and the risk of coronary heart disease in the Normative Aging Study. *American Journal of Cardiology*, 82(7), 851-856.

Shin, D., Oh, Y. H., Eom, C. S. & Park, S. M. (2014) Use of selective serotonin reuptake inhibitors and risk of stroke: a systematic review and meta-analysis. *Journal of Neurology*, 261(4), 686-695.

Singh-Manoux, A., Dugravot, A., Fournier, A., Abell, J., Ebmeier, K., Kivimäki, M. & Sabia, S. (2017) Trajectories of depressive symptoms before diagnosis of dementia: A 28-year follow-up study. *JAMA Psychiatry*, 74(7), 712-718.

Sjösten, N., Nabi, H., Westerlund, H., Salo, P., Oksanen, T., Pentti, J., Virtanen, M., Kivimäki, M. & Vahtera, J. (2013) Effect of depression onset on adherence to medication among hypertensive patients: a longitudinal modelling study. *Journal of Hypertension*, 31(7), 1477-1484.

Skorikov, V. B. & Vandervoort, D. J. (2003) Relationships between the underlying constructs of the Beck Depression Inventory and the Center for Epidemiological Studies Depression Scale. *Educational and Psychological Measurement*, 63(2), 319-335.

Smith, D. J., Nicholl, B. I., Cullen, B., Martin, D., Ul-Haq, Z., Evans, J., Gill, J. M. R., Roberts, B., Gallacher, J., Mackay, D., Hotopf, M., Deary, I., Craddock, N. & Pell, J. P. (2013) Prevalence and Characteristics of Probable Major Depression and Bipolar Disorder within UK Biobank: Cross-Sectional Study of 172,751 Participants. *PLOS ONE*, 8(11), e75362.

Snaith, R. P. (2002) The Hospital Anxiety and Depression Scale. *Journal Of Psychosomatic Research*, 52(5), 401-401.

Solmi, F., Dykxhoorn, J. & Kirkbride, J. B. (2017) Urban-Rural Differences in Major Mental Health Conditions, in Okkels, N., Kristiansen, C. B. & Munk-Jørgensen, P. (eds), *Mental Health and Illness in the City*. Singapore: Springer Singapore, 27-132.

Stansfeld, S. A., Fuhrer, R., Shipley, M. J. & Marmot, M. G. (2002) Psychological distress as a risk factor for coronary heart disease in the Whitehall II Study. *International Journal of Epidemiology*, 31(1), 248-255.

Stapelberg, N. J. C., Neumann, D. L., Shum, D., H. K., McConnell, H. & Hamilton-Craig, I. (2013) The Impact of Methodology and Confounding Variables on the Association Between Major Depression and Coronary Heart Disease: Review and Recommendations. *Current Psychiatry Reviews*, 9(4), 342-352.

Stetler, C. & Miller, G. E. (2011) Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosomatic Medicine*, 73(2), 114-126.

Stewart, J. C., Hawkins, M. A., Khambaty, T., Perkins, A. J. & Callahan, C. M. (2016) Depression and anxiety screens as predictors of 8-year incidence of myocardial infarction and stroke in primary care patients. *Psychosomatic Medicine*, 78(5), 593-601.

Stewart, J. C., Perkins, A. J. & Callahan, C. M. (2014) Effect of collaborative care for depression on risk of cardiovascular events: data from the IMPACT randomized controlled trial. *Psychosomatic Medicine*, 76(1), 29-37.

Stewart, L., Moher, D. & Shekelle, P. (2012) Why prospective registration of systematic reviews makes sense. *Systematic Reviews*, 1(1), 7.

Stürmer, T., Hasselbach, P. & Amelang, M. (2006) Personality, lifestyle, and risk of cardiovascular disease and cancer: follow-up of population based cohort. *BMJ*, 332, 1359.

Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T. & Collins, R. (2015) UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Medicine*, 12(3), e1001779.

Sullivan, P. F., Daly, M. J. & O'Donovan, M. (2012) Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature Reviews Genetics*, 13(8), 537-551.

Sun, J., Ma, H., Yu, C., Lv, J., Guo, Y., Bian, Z., Yang, L., Chen, Y., Shen, H., Chen, Z., Hu, Z. & Li, L. (2016) Association of Major Depressive Episodes with Stroke Risk in a Prospective Study of 0.5 Million Chinese Adults. *Stroke*, 47(9), 2203-2208.

Surtees, P. G., Wainwright, N. W., Luben, R. N., Wareham, N. J., Bingham, S. A. & Khaw, K.-T. (2008) Depression and ischemic heart disease mortality: evidence from the EPIC-Norfolk United Kingdom prospective cohort study. *American Journal of Psychiatry*, 165(4), 515-523.

Taylor, W. D., Aizenstein, H. J. & Alexopoulos, G. S. (2013) The vascular depression hypothesis: mechanisms linking vascular disease with depression. *Molecular Psychiatry*, 18(9), 963-974.

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Andrade, L., Borges, G., Bruffaerts, R., Bunting, B., de Almeida, J. M., Florescu, S., de Girolamo, G., Gureje, O., Haro, J. M., He, Y., Hinkov, H., Karam, E., Kawakami, N., Lee, S., Navarro-Mateu, F., Piazza, M., Posada-Villa, J., de Galvis, Y. T. & Kessler, R. C. (2017) Undertreatment of people with major depressive disorder in 21 countries. *British Journal of Psychiatry*, 210(2), 119-124.

Thorp, J. G., Marees, A., Ong, J.-S., An, J., MacGregor, S. & Derks, E. (2019) Investigating Genetic Heterogeneity in Major Depression Through Item-level Genetic Analyses of the PHQ-9. *bioRxiv*, 528067.

Thygesen, K., Alpert, J. S., Jaffe, A. S., Simoons, M. L., Chaitman, B. R. & White, H. D. (2012) Third Universal Definition of Myocardial Infarction. *Circulation*, 126(16), 2020-2035.

Towfighi, A., Ovbiagele, B., El Husseini, N., Hackett, M. L., Jorge, R. E., Kissela, B. M., Mitchell, P. H., Skolarus, L. E., Whooley, M. A. & Williams, L. S. (2017) Poststroke Depression: A Scientific Statement for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke*, 48(2), e30-e43.

Townsend, P. (2009) Deprivation. *Journal of Social Policy*, 16(2), 125-146.

Trajkova, S., d'Errico, A., Soffietti, R., Sacerdote, C. & Ricceri, F. (2019) Use of Antidepressants and Risk of Incident Stroke: A Systematic Review and Meta-Analysis. *Neuroepidemiology*.

Tran, T., Hammarberg, K., Ryan, J., Lowthian, J., Freak-Poli, R., Owen, A., Kirkman, M., Curtis, A., Rowe, H., Brown, H., Ward, S., Britt, C. & Fisher, J. (2019) Mental health trajectories among women in Australia as they age. *Aging & Mental Health*, 23(7), 887-896.

UK Biobank (2018) *Data providers and dates of data availability*

UK Biobank Data Management Team (2017) *Townsend Deprivation Score* 10/04/2017.

UK Chief Medical Officer (2016) *UK Chief Medical Officer's Low Risk Drinking Guidelines* Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/545937/UK_CMOs__report.pdf [Accessed 28/08/2018].

Undela, K., Parthasarathi, G. & John, S. S. (2015) Impact of antidepressants use on risk of myocardial infarction: A systematic review and meta-analysis. *Indian Journal of Pharmacology*, 47(3), 256-262.

University College London (2019) *Whitehall II - Data Collection.* London: Available online: https://www.ucl.ac.uk/iehc/research/epidemiology-public-health/research/whitehallII/data-collection [Accessed 06/06/2019].

van Buuren, S. (2012) *Flexible imputation of missing data*. Boca Raton, FL, USA: Chapman and Hall/ CRC Press.

van Buuren, S. & Groothuis-Oudshoorn, K. (2011) mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3), 1-67.

van de Schoot, R., Sijbrandij, M., Winter, S. D., Depaoli, S. & Vermunt, J. K. (2017) The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies. *Structural Equation Modeling: A Multidisciplinary Journal*, 24(3), 451-467.

Van der Laan, M. J., Haight, T. J. & Tager, I. B. (2005) Response to "Hypothetical Interventions to Define Causal Effects". *American Journal of Epidemiology*, 162(7), 621-622.

van Eeden, W. A., van Hemert, A. M., Carlier, I. V. E., Penninx, B. W. & Giltay, E. J. (2019) Severity, course trajectory, and within-person variability of individual symptoms in patients with major depressive disorder. *Acta Psychiatrica Scandinavica*, 139(2), 194-205.

Visscher, P. M., Hill, W. G. & Wray, N. R. (2008) Heritability in the genomics era — concepts and misconceptions. *Nature Reviews Genetics*, 9(4), 255-266.

Vistisen, D. & Færch, K. (2014) Eighteen year weight trajectories and metabolic markers of diabetes in modernising China: which timescale is most relevant? *Diabetologia*, 57(12), 2605-2606.

Vistisen, D., Witte, D. R., Tabák, A. G., Herder, C., Brunner, E. J., Kivimäki, M. & Færch, K. (2014) Patterns of Obesity Development before the Diagnosis of Type 2 Diabetes: The Whitehall II Cohort Study. *PLOS Medicine*, 11(2), e1001602.

Wahlbeck, K., Westman, J., Nordentoft, M., Gissler, M. & Laursen, T. M. (2011) Outcomes of Nordic mental health systems: life expectancy of patients with mental disorders. *British Journal of Psychiatry*, 199(6), 453-458.

Wassertheil-Smoller, S., Qi, Q., Dave, T., Mitchell, B. D., Jackson, R. D., Liu, S., Park, K., Salinas, J., Dunn, E. C. & Leira, E. C. (2018) Polygenic Risk for Depression Increases Risk of Ischemic Stroke: From the Stroke Genetics Network Study. *Stroke*, 49(3), 543-548.

Wassertheil-Smoller, S., Shumaker, S., Ockene, J., Talavera, G. A., Greenland, P., Cochrane, B., Robbins, J., Aragaki, A. & Dunbar-Jacob, J. (2004) Depression and Cardiovascular Sequelae in Postmenopausal Women: The Women's Health Initiative (WHI). *Archives of Internal Medicine*, 164(3), 289-298.

Welton, N. J., Caldwell, D. M., Adamopoulos, E. & Vedhara, K. (2009) Mixed Treatment Comparison Meta-Analysis of Complex Interventions: Psychological Interventions in Coronary Heart Disease. *American Journal of Epidemiology*, 169(9), 1158-1165.

Whalley, B., Thompson, D. R. & Taylor, R. S. (2014) Psychological Interventions for Coronary Heart Disease: Cochrane Systematic Review and Meta-analysis. *International Journal of Behavioral Medicine*, 21(1), 109-121.

Whang, W., Kubzansky, L. D., Kawachi, I., Rexrode, K. M., Kroenke, C. H., Glynn, R. J., Garan, H. & Albert, C. M. (2009) Depression and Risk of Sudden Cardiac Death and Coronary Heart Disease in Women Results From the Nurses' Health Study. *Journal of the American College of Cardiology*, 53(11), 950-958.

Whiteford, H. A., Degenhardt, L., Rehm, J., Baxter, A. J., Ferrari, A. J., Erskine, H. E., Charlson, F. J., Norman, R. E., Flaxman, A. D., Johns, N., Burstein, R., Murray, C. J. L. & Vos, T. (2013) Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet*, 382(9904), 1575-1586.

Whooley, M. A., de Jonge, P., Vittinghoff, E., Otte, C., Moos, R., Carney, R. M., Ali, S., Dowray, S., Na, B., Feldman, M. D., Schiller, N. B. & Browner, W. S. (2008) Depressive symptoms, health behaviors, and risk of cardiovascular events in patients with coronary heart disease. *JAMA*, 300(20), 2379-2388.

Whooley, M. A. & Wong, J. M. (2013) Depression and Cardiovascular Disorders. *Annual Review of Clinical Psychology*, 9, 327-354.

Win, S., Parakh, K., Eze-Nliam, C. M., Gottdiener, J. S., Kop, W. J. & Ziegelstein, R. C. (2011) Depressive symptoms, physical inactivity and risk of cardiovascular mortality in older adults: the Cardiovascular Health Study. *Heart*, 97(6), 500-505.

Wong, J., Motulsky, A., Abrahamowicz, M., Eguale, T., Buckeridge, D. L. & Tamblyn, R. (2017) Off-label indications for antidepressants in primary care: descriptive

study of prescriptions from an indication based electronic prescribing system. *BMJ*, 356, j603.

Wong, J., Motulsky, A., Eguale, T., Buckeridge, D. L., Abrahamowicz, M. & Tamblyn, R. (2016) Treatment indications for antidepressants prescribed in primary care in quebec, canada, 2006-2015. *JAMA*, 315(20), 2230-2232.

Woodfield, R., Grant, I. & Sudlow, C. L. (2015) Accuracy of Electronic Health Record Data for Identifying Stroke Cases in Large-Scale Epidemiological Studies: A Systematic Review from the UK Biobank Stroke Outcomes Group. *PLOS ONE*, 10(10), e0140533.

World Health Organisation (2000) *Obesity: Preventing and managing the global epidemic: report of a WHO consultation*. 894. Geneva, Switzerland: Available online: file:///C:/Users/s1570624/Downloads/WHO_TRS_894.pdf [Accessed 07/11/2018].

World Health Organisation (2015) *Excess Mortality in Severe Mental Disorder*.Geneva.

World Health Organisation (2016a) *International Classification of Diseases.* World Health Organisation. Available online: http://www.who.int/classifications/icd/en/ [Accessed 20/07/2016].

World Health Organisation (2016b) *International Classification of Diseases and Related Health Problems 10th Revision.* Available online: http://apps.who.int/classifications/icd10/browse/2016/en [Accessed 20/07/2016].

World Health Organisation (2017) *Cardiovascular diseases (CVDs).* Available online: http://www.who.int/mediacentre/factsheets/fs317/en/ [Accessed 20/05/2017].

World Health Organisation and Food and Agriculture Organisation of the United Nations (2004) *Fruit and vegetables for health : Report of a Joint FAO/WHO Workshop*.Kobe, Japan.

Wouts, L., Voshaar, R. C. O., Bremmer, M. A., Buitelaar, J. K., Penninx, B. W. J. H. & Beekman, A. T. F. (2008) Cardiac Disease, Depressive Symptoms, and Incident Stroke in an Elderly Population. *Archives of General Psychiatry*, 65(5), 596-602.

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S.-A., Bækvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschøn, H. N., Bybjerg-Grauholm, J., Cai, N., Castelao, E., Christensen, J. H., Clarke, T.-K., Coleman, J. I. R., Colodro-Conde, L., Couvy-Duchesne, B., Craddock, N., Crawford, G. E., Crowley, C. A., Dashti, H. S., Davies, G., Deary, I. J., Degenhardt, F., Derks, E. M., Direk, N., Dolan, C. V., Dunn, E. C., Eley, T. C., Eriksson, N., Escott-Price, V., Kiadeh, F. H. F., Finucane, H. K., Forstner, A. J., Frank, J., Gaspar, H. A., Gill, M., Giusti-Rodríguez, P., Goes, F. S., Gordon, S. D., Grove, J., Hall, L. S., Hannon, E., Hansen, C. S., Hansen, T. F., Herms, S., Hickie, I. B., Hoffmann, P., Homuth, G., Horn, C., Hottenga, J.-J., Hougaard, D. M., Hu, M., Hyde, C. L., Ising, M., Jansen, R., Jin, F., Jorgenson, E., Knowles, J. A., Kohane, I. S., Kraft, J., Kretzschmar, W. W., Krogh, J., Kutalik, Z., Lane, J. M., Li, Y., Li, Y., Lind, P. A., Liu, X., Lu, L., MacIntyre, D. J., MacKinnon, D. F., Maier, R. M., Maier, W., Marchini, J., Mbarek, H., McGrath, P., McGuffin, P., Medland, S. E., Mehta, D., Middeldorp, C. M., Mihailov, E., Milaneschi, Y., Milani, L., Mill, J., Mondimore, F. M., Montgomery, G. W., Mostafavi, S., Mullins, N., Nauck, M., Ng, B., et al (2018)

Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, 50(5), 668-681.

Wu, C. S., Wang, S. C., Cheng, Y. C. & Gau, S. S. (2011) Association of cerebrovascular events with antidepressant use: a case-crossover study. *The American Journal of Psychiatry*, 168(5), 511-521.

Wu, Q. & Kling, J. M. (2016) Depression and the Risk of Myocardial Infarction and Coronary Death: A Meta-Analysis of Prospective Cohort Studies. *Medicine*, 95(6), e2815.

Wulsin, L. R. & Singal, B. M. (2003) Do depressive symptoms increase the risk for the onset of coronary disease? A systematic quantitative review. *Psychosomatic Medicine*, 65(2), 201-210.

Xu, X., Mishra, G. D. & Jones, M. (2019) Trajectories of mental health symptoms for women in their 20s predict the onset of chronic physical conditions in their 30s: Two decades of follow-up of young Australian women. *Journal of Affective Disorders*, 246, 394-400.

Yan, T., Escarce, J. J., Liang, L., Longstreth Jr, W. T., Merkin, S. S., Ovbiagele, B., Vassar, S. D., Seeman, T., Sarkisian, C. & Brown, A. F. (2013) Exploring psychosocial pathways between neighbourhood characteristics and stroke in older adults: The cardiovascular health study. *Age and Ageing*, 42(3), 391-397.

Yoshida, K., Chipman, J. J., Bohn, J., Dagostino McGowan, L., Barrett, M. & Haubo Bojesen Christensen, R. (2019) *Create 'Table 1' to Describe Baseline Characteristics*. Available online: https://github.com/kaz-yos/tableone [Accessed 12/12/2019].

Yu, Z. H., Jiang, H. Y., Shao, L., Zhou, Y. Y., Shi, H. Y. & Ruan, B. (2016) Use of antipsychotics and risk of myocardial infarction: a systematic review and meta-analysis. *British Journal of Clinical Pharmacology*, 82(3), 624-632.

Zbozinek, T. D., Rose, R. D., Wolitzky-Taylor, K. B., Sherbourne, C., Sullivan, G., Stein, M. B., Roy-Byrne, P. P. & Craske, M. G. (2012) Diagnostic overlap of generalized anxiety disorder and major depressive disorder in a primary care sample. *Depression and Anxiety*, 29(12), 1065-1071.

Zimmerman, M., Ellison, W., Young, D., Chelminski, I. & Dalrymple, K. (2015) How many different ways do patients meet the diagnostic criteria for major depressive disorder? *Comprehensive Psychiatry*, 56, 29-34.

Zivkovic, S., Koh, C. H., Kaza, N. & Jackson, C. A. (2019) Antipsychotic drug use and risk of stroke and myocardial infarction: a systematic review and meta-analysis. *BMC Psychiatry*, 19(1), 189.

Appendices

A.I. Systematic review

Appendix 1: Search strategy - EMBASE

1. exp depression/
2. depressi*.mp.
3. dysthymi*.mp.
4. MDD.mp.
5. 1 or 2 or 3 or 4
6. cerebrovascular disease/
7. exp basal ganglion haemorrhage/
8. exp brain haemorrhage/
9. brain infarction/
10. exp brain ischemia/
11. carotid artery disease/
12. cerebral artery disease/
13. exp cerebrovascular accident/
14. occlusive cerebrovascular disease/
15. brain embolism/
16. stroke.mp.
17. ((brain$ or cerebr$ or intracerebr$ or intracran$ or subarachnoid) adj5 (haemorrhag$ or haemorrhag$ or haematoma$ or hematoma$ or bleed$)).mp.
18. ((brain$ or cerebr$ or intracerebr$ or intracran$) adj5 (ischemi$ or ischaemi$ or infarct$ or thrombo$ or embol$ or occclus$)).mp.
19. 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18
20. exp heart infarction/
21. ((heart or myocardial or cardiac) adj5 (infarct$ or attack)).mp.
22. 20 or 21
23. 19 or 22
24. longitudinal study/
25. prospective study/
26. follow up/
27. incidence/

28. cohort analysis/

29. etiology/

30. pathogenesis/

31. path analysis/

32. structural equation modeling/

33. epidemiology/

34. comorbidity/

35. (causal$ or causat$).mp.

36. ((direct or indirect) adj3 effect).mp.

37. (pathway$ or mechanism$).mp.

38. (structural equation model$ or "path analys$").mp.

39. mental-physical comorbidit*.mp.

40. physical-mental comorbidit*.mp.

41. cohort.mp.

42. prospective.mp.

43. longitudinal.mp.

44. followup.mp.

45. follow-up.mp.

46. 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45

47. 5 and 23 and 46

48. limit 47 to embase

Appendix 2: Search strategy - Medline

1. Depression/

2. exp Depressive Disorder/

3. depressi*.mp.

4. dysthymi*.mp.

5. MDD.mp.

6. 1 or 2 or 3 or 4 or 5

7. Cerebrovascular Disorders/

8. exp Basal Ganglia Cerebrovascular Disease/

9. exp Brain Ischemia/

10. Carotid Artery Diseases/

11. Intracranial Arterial Diseases/

12. exp "Intracranial Embolism and Thrombosis"/

13. exp Intracranial Haemorrhages/

14. exp Stroke/

15. Vasospasm, Intracranial/

16. (stroke or apoplex*).mp.

17. ((brain$ or cerebr$ or intracerebr$ or intracran$ or subarachnoid) adj5 (haemorrhag$ or haemorrhag$ or haematoma$ or hematoma$ or bleed$)).mp.

18. ((brain$ or cerebr$ or intracerebr$ or intracran$) adj5 (ischemi$ or ischaemi$ or infarct$ or thrombo$ or embol$ or occclus$)).mp.

19. 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18

20. exp Myocardial Infarction/

21. ((heart or myocardial or cardiac) adj5 (infarct$ or attack)).mp.

22. 20 or 21

23. 19 or 22

24. exp Cohort Studies/

25. incidence/

26. Epidemiology/

27. exp causality/

28. Comorbidity/

29. (causal$ or causat$).mp.

30. ((direct or indirect) adj3 effect).mp.

31. (pathway$ or mechanism$).mp.

32. (structural equation model$ or "path analys$").mp.

33. mental-physical comorbidit*.mp.

34. physical-mental comorbidit*.mp.

35. cohort.mp.

36. prospective.mp.

37. longitudinal.mp.

38. followup.mp.

39. follow-up.mp.

40. 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39

41. 6 and 23 and 40

Appendix 3: Search strategy - PsycINFO

1. exp Major Depression/
2. "Depression (Emotion)"/
3. depressi*.mp.
4. dysthymi*.mp.
5. MDD.mp.
6. 1 or 2 or 3 or 4 or 5
7. cerebrovascular accidents/
8. exp cerebrovascular disorders/
9. (stroke or apoplex*).mp.
10. ((brain$ or cerebr$ or intracerebr$ or intracran$) adj5 (ischemi$ or ischaemi$ or infarct$ or thrombo$ or embol$ or occclus$)).mp.
11. ((brain$ or cerebr$ or intracerebr$ or intracran$ or subarachnoid) adj5 (haemorrhag$ or haemorrhag$ or haematoma$ or hematoma$ or bleed$)).mp.
12. 7 or 8 or 9 or 10 or 11
13. exp myocardial infarctions/
14. heart disorders/
15. ((heart or myocardial or cardiac) adj5 (infarct$ or attack)).mp.
16. 13 or 14 or 15
17. 12 or 16
18. exp longitudinal studies/
19. followup studies/
20. epidemiology/
21. exp causality/
22. causal analysis/
23. path analysis/
24. structural equation modeling/
25. risk factors/
26. cohort analysis/
27. etiology/
28. comorbidity/

29. (causal$ or causat$).mp.

30. ((direct or indirect) adj3 effect).mp.

31. (pathway$ or mechanism$).mp.

32. (structural equation model$ or "path analys$").mp.

33. mental-physical comorbidit*.mp.

34. physical-mental comorbidit*.mp.

35. cohort.mp.

36. prospective.mp.

37. longitudinal.mp.

38. followup.mp.

39. follow-up.mp.

40. 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39

41. 6 and 17 and 40

Appendix 4: Search strategy - Web of Science

1. #11 AND #8 AND #1

2. TS=(longitudinal OR "follow up" OR follow-up OR followup OR prospective OR cohort)

3. TS=(stroke)

4. TS=((brain$ or cerebr$ or intracerebr$ or intracran$ or subarachnoid) NEAR/5 (haemorrhag$ or haemorrhag$ or haematoma$ or hematoma$ or bleed$))

5. TS=((brain$ or cerebr$ or intracerebr$ or intracran$) NEAR/5 (isch*mi$ or infarct$ or thrombo$ or embol$ or occclus$))

6. #5 OR #4 OR #3

7. #6 AND #2 AND #1

8. TS=(longitudinal OR "follow up" OR follow-up OR followup OR prospective OR cohort OR mediat* OR "structural equation model*" OR causal* OR "mental-physical comorbidit*" OR "physical-mental comorbidity" OR mechanism* OR pathway*)

9. #8 AND #6 AND #1

10. TS=((heart or myocardial or cardiac) NEAR/5 (infarct$ or attack))

11. #10 OR #6

12. #11 AND #8 AND #1

Appendix 5: Search strategy – CINAHL

S1. (MH "Depression") OR (MH "Dysthymic Disorder")

S2. (TX "MDD") OR (TX "depressi*") OR (TX "dysthymi*")

S3. (MH "Cerebrovascular Disorders") OR (MH "Basal Ganglia Cerebrovascular Disease+") OR (MH "Hypoxia-Ischemia, Brain") OR (MH "Cerebral Ischemia+") OR (MH "Cerebral Small Vessel Diseases") OR (MH "Intracranial Arterial Diseases") OR (MH "Intracranial Embolism and Thrombosis+") OR (MH "Carotid Artery Diseases") OR (MH "Stroke+") OR (MH "Intracranial Haemorrhage+") OR (MH "Cerebral Vasospasm")

S4. TX stroke

S5. TX ((brain* or cerebr* or intracerebr* or intracran* or subarachnoid) N5 (haemorrhag* or haemorrhag* or haematoma* or hematoma* or bleed*))

S6. TX ((brain* or cerebr* or intracerebr* or intracran*) N5 (ischemi* or ischaemi* or infarct* or thrombo* or embol* or occclus*))

S7. (MH "Myocardial Infarction+")

S8. TX ((heart or myocardial or cardiac) N5 (infarct* or attack))

S9. (MH "Causal Modeling+") OR (MH "Comorbidity") OR (MH "Incidence") OR (MH "Postexposure Follow-Up") OR (MH "Prospective Studies+")

S10. (TX "prospective") OR (TX "longitudinal") OR (TX "cohort") OR (TX "follow-up") OR (TX "follow up") OR (TX "followup") OR (TX "mediat*") OR (TX "structural equation model*") OR (TX "causal*") OR (TX "mental-physical comorbidit*") OR (TX "physical-mental comorbidit*") OR (TX "mechanism*") OR (TX "pathway*" OR (TX "causal*) OR (TX "causat*) OR (TX "pathway*) OR (TX "mechanism*") OR (TX "structural equation model*") OR (TX "path analys*) OR (TX ((direct or indirect) N3 effect))

S11. S1 OR S2

S12. S3 OR S4 OR S5 OR S6 OR S7 OR S8

S13. S9 OR S10

S14. S11 AND S12 AND S13

Appendix 6: Example of completed CASP Cohort Study Checklist

(A) Are the results of the study valid?

1. Did the study address a clearly focussed issue?

   Yes – to examine the relationship between depressive symptoms and subsequent ischaemic stroke & to assess the potential mediating role of inflammation

2. Was the cohort recruited in an acceptable way?

   Mostly – Random sample of Medicare-eligible, non-institutionalised adults, aged ≥65 years, residing in one of four US counties, recruited 1989 – 1991. A second enrolment of African-Americans completed 1992 – 1993. Participants being wheelchair bound in the home, undergoing radiation or chemotherapy for cancer and with a haemorrhagic stroke during follow-up were excluded.

   Exclusion of individuals with haemorrhagic stroke might be problematic.

3. Was the exposure accurately measured to minimise bias?

   No – Validated scale but only assessed at baseline.

   Depressive symptoms at baseline, using the 10 item CES-D rating scale, collected by trained interviewers, no time-updated information

4. Was the outcome accurately measured to minimise bias?

   Can't tell – Self-report led to review of medical records, without self-report no diagnosis:

   Self-report of events every 6 months, then confirmation of events or deaths through review of medical records (death records and hospital records), raters were blinded to exposure status.

5. A) Have the authors identified all important confounding factors?

   Mostly, did not identify physical inactivity, alcohol use, and medications

   B) Have they taken account of the confounding factors in the design and/ or analysis?

   Yes – Adjustment in statistical analyses, stratification by c-reactive protein tertiles

   C) Have the authors discussed mechanisms/ mediating factors?

   Yes – inflammation (c-reactive protein)

   D) Have they taken account of the mediating factors in the design and/ or analysis?

   Yes – taken into account by stratifying the analysis

6. A) Was the follow-up of subjects complete enough?

   Can't tell – Not reported

B) Was the follow-up of subjects long enough?

Yes – Median follow-up: 11 years but no range reported

(B) What are the results of this study?

7. What are the results of this study?

   No association between depressive symptoms and haemorrhagic stroke (participants with haemorrhagic stroke were excluded due to this), positive association between depressive symptoms and ischaemic stroke, j-shaped relationship with c-reactive protein tertiles, authors concluded that it did not act as mediator (investigated by adjusting for it in cox model), results were similar for men and women

8. How precise are the results?

   Relatively precise (HR: 1.25, 95% CI: 1.02 – 1.53 in fully adjusted model)

9. Do you believe the results?

   + Stroke at baseline excluded

   + Attempted to investigate mediation but…

   - Used adjustment in cox model to investigate mediation

   - Participants with haemorrhagic stroke excluded (potential selection bias)

   - Small sample size within tertiles

(C) Will the results help locally?

10. Can the results be applied to the local population?

    Applicable to Medicare-insured adults but selection bias might explain some of the risk. Difficult to know to what extent sample sufficiently differs from local population to cause concern

11. Do the results of this study fit with other available evidence?

    Yes – all included studies on the association between depressive symptoms and risk of ischaemic stroke showed an increased risk among individuals with high depressive symptoms, relative to non-exposed individuals

12. What are the implications of this study for practice?

    Results of one observational study not sufficiently robust to recommend changes to clinical practice or health policy decision making

Table A 1: Data sources used to determine events during follow-up (page 1 of 4)

Author(s), year	MI as outcome?	Stroke as outcome?	Self-report	Proxy report	Physician report/ confirmation	GP records	Hospital records	Medical records*	Death records	Coroner/ autopsy reports	Others
Arbelaez et al (2007)		Ischaemic stroke	✓	✓			✓	✓	✓		
Ariyo et al (2000)	MI		✓		✓		✓	✓	✓	✓	Obituaries
Avendano et al (2006)		Total stroke	✓	✓			✓		✓		Obituaries
Barefoot & Schroll (1996)	MI						✓		✓	✓	
Bos et al (2008)		Total stroke, ischaemic stroke				✓	✓				Municipality records
Brown et al (2011)	MI							✓†			
Brunner et al (2014)		Total stroke			✓		✓		✓		
Chi et al (2014)	MI										Insurance records
Cummings et al (2016)	MI	Total stroke	✓	✓			✓		✓	✓	
Daskalopoulou et al (2016)	MI	Intracerebral haemorrhage, subarachnoid haemorrhage, ischaemic stroke			✓		✓		✓‡		The Myocardial Ischaemia National Audit Project§
Everson et al (1998)		Total stroke							✓		
Everson-Rose et al (2014)		Total stroke, ischaemic stroke	✓	✓	✓		✓	✓	✓		
Gafarov et al (2017)	MI	Total stroke	✓‡	✓‡			✓‡	✓‡	✓‡		WHO Registry§
Gilsanz et al (2015)		Total stroke	✓	✓							
Gilsanz et al (2017)		Total stroke, ischaemic stroke, haemorrhagic stroke	✓	✓			✓		✓		
Glymour et al (2010)		Total stroke	✓	✓							
Glymour et al (2012)		Total stroke	✓	✓							
Gustad et al (2014)	MI						✓		✓		

Table A1 continued: Data sources used to determine events during follow-up (page 2 of 4)

Author(s), year	MI as outcome?	Stroke as outcome?	Self-report	Proxy report	Physician report/ confirmation	GP records	Hospital records	Medical records*	Death records	Coroner/ autopsy reports	Others
Henderson et al (2013)		Haemorrhagic stroke					✓				
Jackson & Mishra (2013)		Total stroke	✓						✓		
Janszky et al (2010)	MI						✓		✓		
Joyce (2015)	MI						✓				
Kim et al (2011)		Total stroke	✓	✓							
Kim et al (2013)		Total stroke	✓	✓							
Köhler et al (2013)		Total stroke			✓						
Kubzansky et al (2006)	MI						✓				
Langvik & Hjemdal (2015)	MI		✓								
Larson et al (2001)		Total stroke	✓						✓		
Lin et al (2014)	MI										Insurance records
Majed et al (2012)		Total stroke, ischaemic stroke	✓			✓	✓				
Marijnissen et al (2014)		Total stroke	✓		✓				✓		
Mathur et al (2016)	MI	Total stroke				✓					
Mejía-Lancheros et al (2014)	MI	Total stroke	✓	✓		✓		✓	✓		
Moise et al (2016)		Total stroke	✓	✓		✓	✓		✓	✓	
Niles & O'Donovan (2018)		Total stroke	✓								
O'Brien et al (2015)		Total stroke	✓				✓		✓		
Ohira et al (2001)		Total stroke, ischaemic stroke, haemorrhagic stroke	✓	✓	✓			✓	✓		Ambulance records, reports by public health nurses and health volunteers

Table A1 continued: Data sources used to determine events during follow-up (page 3 of 4)

Author(s), year	MI as outcome?	Stroke as outcome?	Self-report	Proxy report	Physician report/ confirmation	GP records	Hospital records	Medical records*	Death records	Coroner/ autopsy reports	Others
Pan et al (2011a)		Total stroke, ischaemic stroke, haemorrhagic stroke, stroke of unknown type	✓	✓				✓	✓	✓	Postal authorities
Péquignot et al (2013)	MI	Total stroke	✓	✓	✓		✓		✓	✓	
Péquignot et al (2016)		Total stroke	✓	✓	✓		✓		✓	✓	Emergency records
Pössel et al (2015)	MI						✓				
Pratt et al (1996)	MI		✓								
Scherrer et al (2010)	MI								✓**		
Scherrer et al (2011)	MI								✓**		
Sesso et al (1998)††	MI										
Stewart et al (2016)	MI	Total stroke					✓	✓†	✓		
Stürmer et al (2006)	MI	Total stroke	✓		✓				✓		
Sun et al (2016)		Total stroke, ischaemic stroke, haemorrhagic stroke						✓	✓		
Wassertheil-Smoller et al (2004)	MI	Total stroke	✓	✓				✓	✓		
Wouts et al (2008)		Total stroke	✓		✓				✓		
Yan et al (2013)		Ischaemic stroke	✓				✓				

[*] Not further specified unless indicated otherwise
[†] Specified as Regenstrief Medical Record System but not further specified what records are collected as part of this system
[‡] Stroke only
[§] MI only
[**] Specified as Veterans' Health Administration records (inpatient and outpatient) but not further specified what records are collected as part of this system
[††] Data sources not specified
GP: General practitioner, MI: Myocardial infarction

Table A 2: Results of studies on the risk of total stroke among individuals with clinical depression or depressive symptoms, relative to non-exposed individuals (page 1 of 6)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Avendano et al (2006), aged ≥ 75 years	Depressive symptoms (CES-D ≥ 21)	104 / 958	NR	NR	HR		0.95 (0.46 – 1.98)
Avendano et al (2006), aged 65 – 74 years		122 / 1,292					3.05 (1.63 – 5.70)
Bos et al (2008)	Depressive symptoms (CES-D ≥ 16)	291 / 4,424	NR / 324	NR / 4,100	HR	1.20 (0.81 – 1.80)‡	1.21 (0.80 – 1.83)
Brunner et al (2014)	Depressive symptoms (CES-D ≥ 16)	58 / 5,717	8 / 845	50 / 4,872	HR		1.21 (0.61 – 2.42)
Cummings et al (2016), no diabetes	Depressive symptoms (CES-D-4 ≥ 4) and elevated stress (CPSS > 4) versus neither depressive symptoms nor elevated stress	362 / 13,738	24 / 1,118	257 / 12,620	HR	1.15 (0.76 – 1.74)	1.14 (0.75 – 1.75)
Cummings et al (2016), diabetes		124 / 4,090	14 / 416	68 / 2,583		1.40 (0.79 – 2.49)	1.57 (0.86 – 2.87)
Everson et al (1998)	Depressive symptoms (HPL ≥ 5)	169 / 6,676	39 / 969	130 / 5,707	HR	1.94 (1.36 – 2.78)	1.54 (1.06 – 2.22)
	Time-varying depressive symptoms (HPL ≥ 5)	101 / NR	NR	NR			1.55 (0.97 – 2.47)
Everson-Rose et al (2014)	Depressive symptoms (CES-D: Quartiles were created, top group was further split into two groups: highest versus lowest group)	144 / 6,643	26 / 853	36 / 1,805	HR		1.76 (1.03 – 2.99)
Gafarov et al (2017), men	Depressive symptoms (MOPSY, no cut-off reported)	22 / 190	NR	NR	HR	5.84 (2.47 – 13.76)	4.20 (1.60 – 11.10)
Gafarov et al (2017), women		35 / 384				4.63 (1.03 – 20.90)	8.50 (1.06 – 69.30)

Table A 2 continued: Results of studies on the risk of total stroke among individuals with clinical depression or depressive symptoms, relative to non-exposed individuals (page 2 of 6)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Gilsanz et al (2015)	Recent onset depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms	1,192 / 55,731 person observations	NR / 5,373 person observations	NR / 39,080 person observations	HR		1.08 (0.81 – 1.44)
	Stable high depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 6,111 person observations				2.14 (1.69 – 2.71)
	Recently remitted depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 5,167 person observations				1.66 (1.22 – 2.26)
Gilsanz et al (2017)	Recent onset depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms	334 / 24,940 person observations	NR / 1,499 person observations	NR / 19,552 person observations	HR		1.44 (0.97 – 2.14)
	Stable high depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 1,893 person observations				1.65 (1.06 – 2.56)
	Recently remitted depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 2,026 person observations				1.02 (0.66 – 1.58)
Glymour et al (2010)	Depressive symptoms (8 item CES-D ≥ 3)	1,864 / 19,087	512 / 3,842	1,352 / 15,245	HR		1.23 (1.11 – 1.38)
	Time-varying depressive symptoms (8 item CES-D ≥ 3)						1.53 (1.39 – 1.69)

Table A 2 continued: Results of studies on the risk of total stroke among individuals with clinical depression or depressive symptoms, relative to non-exposed individuals (page 3 of 6)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Glymour et al (2012), Hispanic adults	Depressive symptoms (8 item CES-D ≥ 3)	121 / 1,313	50 / 433	71 / 880	HR		0.99 (0.66 – 1.5)
Glymour et al (2012), White/ other adults		1,543 / 14,778	369 / 2,601	1,174 / 12,177			1.28 (1.13 – 1.44)
Glymour et al (2012), African-American adults		334 / 2,557	111 / 692	223 / 1,865			1.12 (0.88 – 1.42)
Jackson & Mishra (2013)	Time-varying depressive symptoms (10 item CES-D ≥ 10)	143 / 10,547	NR / 2,612	NR / 7,787	OR with GEE		1.96 (1.37 – 2.81)
	Time-varying depressive symptoms (10 item CES-D ≥ 10) or antidepressant use (self-report)					2.41 (1.78 – 3.27)	1.94 (1.37 – 2.74)
	Time-varying depressive symptoms (10 item CES-D ≥ 10) or clinical depression (self-report)						2.1 (1.49 – 2.96)
	Time-varying depressive symptoms (10 item CES-D ≥ 10) or clinical depression (self-report) or antidepressant use (self-report)						2.03 (1.44 – 2.86)
Kim et al (2011)	Depressive symptoms (CES-D, not further specified)	88 / 6,044	NR	NR	OR		1.20 (1.06 – 1.34)
Kim et al (2013)	Depressive symptoms (CES-D, not further specified)	265 / 6,739	NR	NR	OR		1.16 (1.10 – 1.24)
Köhler et al (2013)	Depressive symptoms (GDS-15 ≥ 6)	209 / 2,854	19 / 261	190 / 2,593	HR	1.16 (0.73 – 1.86)	0.90 (0.55 – 1.48)
Larson et al (2001)	Clinical depression (DIS)	95 / 1,703	7 / 101	87 / 1,575	RR		2.67 (1.08 – 6.63)

Table A 2 continued: Results of studies on the risk of total stroke among individuals with clinical depression or depressive symptoms, relative to non-exposed individuals (page 4 of 6)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Majed et al (2012)	Depressive symptoms (13 item CES-D: Highest quartile versus quartile 1 to 3)	125 / 8,746	36 / 6,767	89 / 1,979	HR	1.60 (1.09 – 2.36)	1.41 (0.95 – 2.11)
Marijnissen et al (2014), no cardiac disease	Depressive symptoms (CES-D ≥ 16)	85 / 1,649	NR	NR	HR		42.60 (5.23 – 347)
Marijnissen et al (2014), cardiac disease	Depressive symptoms (CES-D ≥ 16)	32 / 410	NR	NR			0.37 (0.01 – 26.3)
Mathur et al (2016)	Clinical Depression (Read codes)	987 / 524,952	79 / 21,811	908 / 503,141	HR		1.29 (1.00 – 1.66)
Mejía-Lancheros et al (2014)	Clinical Depression (self-report)	136 / 7,263	15 / 1,309	121 / 5,954	HR		0.66 (0.38 – 1.15)
Moise et al (2016)	Time-varying depressive symptoms (4 item CES-D ≥ 4)	663 / 22,666	81 / 2,267	582 / 20,399	HR	1.35 (1.07 – 1.70)	1.26 (0.99 – 1.60)
O'Brien et al (2015)	Depressive symptoms (CES-D ≥ 16)	79 / 3,309	23 / 738	56 / 2,571	HR	1.47 (0.90 – 2.38)	1.28 (0.74 – 2.23)
	Depressive symptoms (CES-D ≥ 21 vs CES-D < 16)		NR / 384	NR / 2,571		2.00 (1.15 – 3.48)	1.95 (1.02 – 3.71)
Ohira et al (2001)	Depressive symptoms (Zung SDS ≥ 35 vs ≤ 30)	69 / 884	30 / 295	19 / 332	HR	2.00 (1.10 – 3.60)‡	1.90 (1.10 – 3.50)
	Depressive symptoms (Zung SDS ≥ 40 vs ≤ 30)		NR	NR			3.50 (1.80 – 6.90)
Pan et al (2011a)	Time-varying current clinical depression and/ or antidepressant use compared to never clinically depression (self-report)	1,033 / 458,463 py	160 / 52,746 py	796 / 376,719 py	HR	1.63 (1.38 – 1.94)‡	1.41 (1.18 – 1.67)
	Time-varying past clinical depression and/ or antidepressant use compared to never clinically depressed (self-report)		77 / 28,998 py			1.40 (1.11 – 1.77)‡	1.23 (0.97 – 1.56)

Table A 2 continued: Results of studies on the risk of total stroke among individuals with clinical depression or depressive symptoms, relative to non-exposed individuals (page 5 of 6)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Péquignot et al (2013)	Depressive symptoms (CES-D ≥ 16)	141 / 7,308	NR / 1,657	NR / 5,651	HR		1.54 (1.06 – 2.25)
		116 / 7,308					1.36 (0.89 – 2.09)
		25 / 7,308					3.27 (1.42 – 7.52)
Péquignot et al (2016)	Depressive symptoms (CES-D ≥ 16 in wave preceding stroke event)	245 / 7,313	NR / 1,657	NR / 5,656	HR		1.48 (1.12 – 1.97)
	Number of waves with depressive symptoms (CES-D ≥ 16, per additional wave)		NA	NA			1.16 (1.03 – 1.31)
Stewart et al (2016)	Depressive symptoms (modified PRIME-MD: either of two items vs none)	235 / 2,041	36 / 270	198 / 1,771	HR		1.13 (0.78 – 1.64)
Stürmer et al (2006)	Factor analysis of personality scales (highest vs. lowest third)	62 / 33,247 py	32 / 11,552 py	12 / 10,329 py	HR	1.74 (0.98 – 3.10)‡	1.53 (0.83 – 2.80)
Sun et al (2016)	Clinical depression (CIDI-SF: MDE within the past year)	27,623 / 487,377	183 / 2,988	27,440 / 484,389	HR	1.14 (0.98 – 1.31)‡	1.14 (0.99 – 1.32)
	Number of depressive symptoms (CIDI-SF: 7 vs 0 – 2)		31 / 450			1.47 (1.04 – 2.08)	1.47 (1.06 – 2.08)
Sun et al (2016)	Depressive symptoms (6 item CES-D ≥ 5) and responded "yes" to 2 DIS item	464 / 73,098	NR	NR	HR		1.01 (0.78 – 1.30)
Wouts et al (2008), no cardiac disease	Clinical depression (DIS) or depressive symptoms (CES-D ≥ 16)	NR / 2,354	NR	NR	HR		0.73 (0.41 – 1.30)
	Clinical depression (DIS)						0.44 (0.06 – 3.22)
Wouts et al (2008), cardiac disease	Clinical depression (DIS) or depressive symptoms (CES-D ≥ 16)	NR / 611					2.18 (1.17 – 4.09)
	Clinical depression (DIS)						2.66 (0.61 – 11.56)

* Data are presented as estimate (95% confidence interval)
† Unless stated otherwise
‡ Age and sex adjusted
CES-D: Centre for Epidemiologic Studies Depression Scale, CIDI-SF: Composite International Diagnostic Interview, CPSS: Cohen Perceived Stress Scale, DIS: Diagnostic Interview Scale, GDS: Geriatric Depression Scale, GEE: Generalised Estimating Equations, HPL: Human Population Laboratory Depression scale, HR: Hazard ratio, MDE: Major depressive episode, MOPSY: Monitoring trends and determinants in cardiovascular disease project psychosocial Interview Depression scale, NA: Not applicable, NR: Not reported, OR: Odds ratio, PRIME-MD: Primary Care Evaluation of Mental Disorders, Py: Person years, RR: Risk ratio, Zung SDS: Zung Self-Rated Depression Scale

Table A 3: Results of studies on the risk of total stroke per one unit increase on depressive symptoms rating scales

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted)	Effect estimate* (most adjusted)
Everson et al (1998)	Depressive symptoms (HPL, 1-point increase)	169 / 6,676	NA	NA		1.09 (1.03 – 1.15)	
	Time-varying depressive symptoms (HPL, 1-point increase)	101 / NR				1.06 (0.99 – 1.14)	
Glymour et al (2010)	Depressive symptoms (8 item CES-D, 1-point increase)	1,864 / 19,087	NA	NA	HR		1.07 (1.05 – 1.1)
Marijnissen et al (2014), no cardiac disease	Depressive symptoms (CES-D, 1-point increase)	85 / 1,649	NA	NA	HR		1.12 (1.03 – 1.22)
Marijnissen et al (2014), cardiac disease		32 / 410					0.97 (0.79 – 1.20)
Niles & O'Donovan (2018)	Depressive symptoms (CES-D, not further specified)	NR / 10,843	NA	NA	OR		1.12 (1.01 – 1.24)
O'Brien et al (2015)	Depressive symptoms (CES-D, 1 SD increase)	79 / 3,309	NA	NA	HR	1.26 (1.04 – 1.53)	1.30 (1.02 – 1.66)
Ohira et al (2001)	Depressive symptoms (Zung SDS, 1 SD/ six points increase)	69 / 884	NA	NA			1.40 (1.10 – 1.70)
Stürmer et al (2006)	Factor analysis of personality scales (1 SD increase)	62 / 33,247 py	NA	NA	HR	1.26 (1.02 – 1.55)	1.13 (0.88 – 1.46)
Wouts et al (2008), no cardiac disease	Depressive symptoms (CES-D, 1-point increase)	NR / 2,354	NA	NA	HR		0.99 (0.96 – 1.01)
Wouts et al (2008), cardiac disease		NR / 611					1.05 (1.02 – 1.08)

* Data are presented as estimate (95% confidence interval)

CES-D: Centre for Epidemiologic Studies Depression Scale, HPL: Human Population Laboratory Depression scale, HR: Hazard ratio, NA: Not applicable, NR: Not reported, OR: Odds ratio, Py: Person years, SD: Standard deviation, Zung SDS: Zung Self-Rated Depression Scale

Table A 4: Results of studies on the risk of ischaemic and uncertain type of stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals (page 1 of 3)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Arbelaez et al (2007)	Depressive symptoms (10 item CES-D ≥ 8)	607 / 5,525	14.9 per 1,000 py	11.4 per 1,000 py	HR	1.29 (1.07 – 1.56)	1.25 (1.02 – 1.53)
Bos et al (2008)	Depressive symptoms (CES-D ≥ 16)	190 / 4,424	NR / 324	NR / 4,100	HR	1.43 (0.89 – 2.31)‡	1.43 (0.87 – 2.35)
Daskalopoulou et al (2016)	History of clinical depression and/ or antidepressant use (CPRD record)	6,053 / 1,937,360	1,312 / 367,117	4,741 / 1,570,243	HR		1.26 (1.17 – 1.34)
	New onset clinical depression (past year) and/ or antidepressant use (CPRD record)	4,287 / 1,356,578	119 / 39,747	4,168 / 1,316,831			1.17 (0.97 – 1.41)
Everson-Rose et al (2014)	Depressive symptoms (CES-D ≥ 16 versus 0 – 2)	120 / 6,643	23 / 853	31 / 1,805	HR		1.89 (1.07 – 3.34)
Gilsanz et al (2017)	Recent onset depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms	285 / 24,924 person observations	NR / 1,499 person observations	NR / 19,552 person observations	HR		1.49 (0.97 – 2.30)
	Stable high depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 1,893 person observations				1.64 (1.04 – 2.60)
	Recently remitted depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 2,026 person observations				1.08 (0.69 – 1.70)
Majed et al (2012)	Depressive symptoms (13 item CES-D, highest quartile versus quartile 1 to 3)	98 / 8,746	32 / 6,767	66 / 1,979	HR	1.92 (1.26 – 2.93)	1.65 (1.07 – 2.55)

Table A 4 continued: Results of studies on the risk of ischaemic and uncertain type of stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals (page 2 of 3)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Ohira et al (2001)	Depressive symptoms (Zung SDS ≥ 35 vs ≤ 30)	39 / 884	20 / 295	9 / 332	HR	3.0 (1.3 – 6.6)	2.7 (1.2 – 6.0)
	Depressive symptoms (Zung SDS ≥ 40 vs ≤ 30)		NR	NR			6.4 (2.5 – 16.1)
	Depressive symptoms (Zung SDS per 1 SD/ six points increase)		NA	NA			1.8 (1.3 – 2.5)
Pan et al (2011a), ischaemic stroke	Time-varying current clinical depression and/ or antidepressant use compared to never depression (self-report)	538 / 458,891 py	79 / 52,817 py	428 / 377,043 py	HR	1.51 (1.18 – 1.91)‡	1.28 (1.00 – 1.63)
	Time-varying past clinical depression and/ or antidepressant use compared to never depression (self-report)		31 / 29,031 py			1.05 (0.73 – 1.52)‡	0.92 (0.64 – 1.33)
Pan et al (2011a), uncertain type of stroke	Time-varying current clinical depression and/ or antidepressant use compared to never depression (self-report)	371 / 459,110 py	64 / 52,836 py	270 / 377,231 py		1.93 (1.47 – 2.53)‡	1.66 (1.26 – 2.19)
	Time-varying past clinical depression and/ or antidepressant use compared to never depression (self-report)		37 / 29,043 py			1.98 (1.40 – 2.80)‡	1.72 (1.21 – 2.43)
Sun et al (2016)	Clinical depression (CIDI-SF: MDE within the past year)	21,427 / 487,377	135 / 2,988	21,292 / 484,389	HR		1.13 (0.96 – 1.34)
Yan et al (2013), White adults	Time-varying depressive symptoms (CES-D, dichotomised, no cut-off reported)	548 / 3,834	116 / 739	432 / 3,095	HR	1.2 (1.01 – 1.52)	1.20 (0.97 – 1.48)
Yan et al (2013), African-American adults		102 / 785	34 / 233	68 / 552		1.29 (0.85 – 1.94)	1.21 (0.79 – 1.84)

* Data are presented as estimate (95% confidence interval)
† Unless stated otherwise
‡ Age and sex adjusted
CES-D: Centre for Epidemiologic Studies Depression Scale, CIDI-SF: Composite International Diagnostic Interview, CPRD: Clinical Practice Research Datalink, HR: Hazard ratio, NA: Not applicable, NR: Not reported, PRIME-MD: Primary Care Evaluation of Mental Disorders, Py: Person years, SD: Standard deviation, Zung SDS: Zung Self-Rated Depression Scale

Table A 5: Results of studies on the risk of haemorrhagic stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals (page 1 of 2)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Daskalopoulou et al (2016), subarachnoid haemorrhage	History of clinical depression and/ or antidepressant use (CPRD record)	1,278 / 1,937,360	315 / 367,117	963 / 1,570,243	HR		1.18 (1.02 – 1.36)
Daskalopoulou et al (2016), intracerebral haemorrhage		2,388 / 1,937,360	1,312 / 367,117	4,741 / 1,570,243			1.31 (1.18 – 1.46)
Daskalopoulou et al (2016), subarachnoid haemorrhage	New onset clinical depression (past year) and/ or antidepressant use (CRRD record) (participants with a history of depression or antidepressant use were excluded)	853 / 1,356,578	34 / 39,747	819 / 1,316,831			1.30 (0.92 – 1.85)
Daskalopoulou et al (2016), intracerebral haemorrhage		1,655 / 1,356,578	45 / 39,747	1,610 / 1,316,831			1.13 (0.84 – 1.52)
Gilsanz et al (2017)	Recent onset depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms	33 / 24,924 person observations	NR / 1,499 person observations	NR / 19,552 person observations	HR		2.16 (0.73 – 6.40
	Stable high depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 1,893 person observations				1.92 (0.66 – 5.54)
	Recently remitted depressive symptoms (8 item CES-D ≥ 3) compared to stable low/ no depressive symptoms		NR / 2,026 person observations				0.45 (0.06 – 3.35)
Henderson et al (2013)	Depressive symptoms (10 item CES-D per 1 SD increase)	44 / 2,539	NA	NA	HR		1.63 ($p < 0.01$)

Table A 5 continued: Results of studies on the risk of haemorrhagic stroke among individuals with depression or depressive symptoms, relative to non-exposed individuals (page 2 of 2)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Ohira et al (2001)	Depressive symptoms (Zung SDS ≥ 35 versus ≤ 30)	20 / 884	5 / 295	6 / 332	HR	0.90 (0.30 – 3.10)	0.90 (0.30 – 3.10)
Pan et al (2011a)	Time-varying current clinical depression and/ or antidepressant use compared to never clinically depression (self-report)	124 / 459,299 py	17 / 52,870 py	98 / 377,371 py	HR	1.39 (0.83 – 2.33)‡	1.28 (0.76 – 2.17)
	Time-varying past clinical depression and/ or antidepressant use compared to never clinically depression (self-report)		9 / 29,058 py			1.32 (0.66 – 2.61)‡	1.22 (0.61 – 2.44)
Sun et al (2016)	Clinical depression (CIDI-SF: MDE within the past year)	5,255 / 487,377	40 / 2,988	5,215 / 484,389	HR		1.16 (0.85 – 1.58)

* Data are presented as estimate (95% confidence interval)
† Unless stated otherwise
‡ Age and sex adjusted
CES-D: Centre for Epidemiologic Studies Depression Scale, CIDI-SF: Composite International Diagnostic Interview, CPRD: Clinical Practice Research Datalink, HR: Hazard ratio, MDE: Major depressive episode, MMPI OBD subscale: Minnesota Multiphasic Personality Inventory - obvious depression subscale, MMPI-2: Second version of the Minnesota Multiphasic Personality Inventory, NA: Not applicable, NR: Not reported, Py: Person years, RR: Risk ratio, SD: Standard deviation, Zung SDS: Zung Self-Rated Depression Scale

Table A 6: Results of studies on the risk of myocardial infarction among individuals with depression or depressive symptoms, relative to non-exposed individuals (page 1 of 3)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Brown et al (2011)	Depressive symptoms (CES-D ≥ 16)	517 / 2,728	103 / 423	415 / 2,305	HR	1.49 (1.20 – 1.85)	1.55 (1.24 – 1.93)
Chi et al (2014)	Clinical depression (ICD-9-CM: 296.2, 296.3, 300.4, 311)	2,146 / 132,090	494 / 26,418	1,652 / 105,672	HR	1.26 (1.14 – 1.39)	1.03 (0.93 – 1.15)
Cummings et al (2016), no diabetes	Neither depressive symptom (CES-D-4 ≥ 4) nor elevated stress (CPSS > 4) versus depressive symptoms and elevated stress	455 / 17,913	28 / 1,118	329 / 12,620	HR	1.27 (0.79 – 2.02)	1.27 (0.86 – 1.88)
Cummings et al (2016), diabetes		203 / 4,090	24 / 416	117 / 2,583		1.39 (0.92 – 2.10)	1.48 (0.96 – 2.27)
Daskalopoulou et al (2016)	History of clinical depression and/ or antidepressant use (CPRD record)	16,239 / 1,937,360	3,257 / 367,117	12,982 / 1,570,243	HR		1.21 (1.16 – 1.27)
Daskalopoulou et al (2016)	New onset clinical depression (past year) and/ or antidepressant use (CPRD record)	11,370 / 1,356,578	285 / 39,747	11,085 / 1,316,831			1.11 (0.99 – 1.3)
Gafarov et al (2017), men	Depressive symptoms (MOPSY, no cut-off reported)	30 / 190	NR	NR	HR	2.00 (1.20 – 3.36)	1.60 (0.90 – 2.80)
Gafarov et al (2017), women		15 / 384				2.53 (1.26 – 24.34)	3.47 (0.51 – 23.68)
Gustad et al (2014)	Depressive symptoms (HADS-D: 0 – 7 versus ≥ 11)	1,561 / 590,365 py	NR	NR	HR	1.32 (1.07 – 1.61)‡	1.31 (1.03 – 1.66)
		1,228 / 51,442 py					1.33 (1.02 – 1.75)
Janszky et al (2010)	Clinical depression (ICD-8: 296 or 300.4)	1,295 / 49,321	22 / 646	1,273 / 48,675	HR	1.36 (0.89 – 2.07)	1.03 (0.65 – 1.65)
Joyce (2015), men	Clinical depression (self-reported doctor diagnosis: No depression & no anxiety versus depression alone)	451 / 60,298	28 / 4,023	380 / 52,240	HR		1.10 (0.75 – 1.62)
Joyce (2015), women		195/ 83,517	19 / 8,855	148 / 65,440			1.20 (0.74 – 1.95)

Table A 6 continued: Results of studies on the risk of myocardial infarction among individuals with depression or depressive symptoms, relative to non-exposed individuals (page 2 of 3)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Kubzansky et al (2006)	Factor analysis identified iso-depression based on MMPI-2 (medium compared to lowest tertile)	70 / 1,306	NR	NR	HR	1.59 (0.7 – 3.8)‡	1.52 (0.6 – 3.6)
Lin et al (2014)	Clinical depression (ICD-9-CM: 296.2, 296.3, 300.4, 311)	941 / 54,355	379 / 10,871	562 / 43,484	HR		1.88 (1.63 – 2.17)
Mathur et al (2016)	Clinical Depression (Read codes)	3,390 / 524,952	263 / 21,811	3,127 / 503,141	HR		1.21 (1.05 – 1.39)
Mejía-Lancheros et al (2014)	Clinical Depression (self-report)	103 / 7,263	13 / 1,309	90 / 5,954	HR		0.89 (0.49 – 1.62)
Péquignot et al (2013)	Depressive symptoms (CES-D ≥ 16)	NR / 7,308	NR / 1,657	NR / 5,651	HR		1.03 (0.63 – 1.7)
Pratt et al (1996)	Clinical depression (DIS: Lifetime history of MDE versus neither dysphoria nor MDE)	64 / 1,551	6 / 73	37 / 1,107	OR	2.59 (1.06 – 6.35)	4.14 (1.48 – 11.62)
Scherrer et al (2010)	Clinical depression (ICD-9-CM: 296.2, 296.3, 300.4 or 311)	12,304 / 355,999	3,961 / 96,612	8,343 / 259,387	HR		1.39 (1.34 – 1.45)
Scherrer et al (2011)	Clinical depression (ICD-9-CM: 296.2, 296.3, or 311)	8,313 / 292,317	2,682 / 77,568	5,631 / 214,749	HR		1.29 (1.22 – 1.37)
Sesso et al (1998)	Depressive symptoms (MMPI-2 D: 4th quartile versus 1st quartile)	30 / 1,305	9 / 315	4 / 329	HR	2.42 (0.74 – 7.88)‡	2.40 (0.74 – 7.85)
	Depressive symptoms (MMPI-2 Dep: 4th quartile versus 1st quartile)		10 / 322	5 / 275		1.86 (0.63 – 5.43)‡	1.75 (0.59 – 5.15)
	Depressive symptoms (SCL-90: 4th quartile versus 1st quartile)		6 / 308	5 / 304		1.27 (0.39 – 4.17)‡	1.27 (0.38 – 4.26)
Stewart et al (2016)	Depressive symptoms (modified PRIME-MD questionnaire: either of two items vs none)	553 / 2,041	77 / 270	476 / 1,771	HR		0.99 (0.77 – 1.28)

Table A 6 continued: Results of studies on the risk of myocardial infarction among individuals with depression or depressive symptoms, relative to non-exposed individuals (page 3 of 3)

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Stürmer et al (2006)	Factor analysis of personality scales (highest vs. lowest third)	72 / 32,875 py	32 / 11,366 py	21 / 10,288 py	HR	0.80 (0.46 – 1.39)	0.80 (0.44 – 1.43)
Wassertheil-Smoller et al (2004)	Depressive symptoms (6 item CES-D ≥ 5) and responded "yes" to two DIS items	509 / 73,098	NR	NR	HR		1.12 (0.89 – 1.41)

* Data are presented as estimate (95% confidence interval)
† Unless stated otherwise
‡ Age and sex adjusted
CES-D: Centre for Epidemiologic Studies Depression Scale, CPRD: Clinical Practice Research Datalink, CPSS: Cohen Perceived Stress Scale, DIS: Diagnostic Interview Scale, HADS-D: Hospital Anxiety and Depression Scale - Depression subscale, HR: Hazard ratio, ICD-8: 8th version of the International Classification of Diseases , ICD-9-CM: Clinical Modification of the 9th version of the International Classification of Diseases, MDE: Major depressive episode, MI: Myocardial infarction, MMPI-2: Second version of the Minnesota Multiphasic Personality Inventory, MMPI-2 D: Second version of the Minnesota Multiphasic Personality Inventory - Depression subscale in keeping with Butcher et al: Development and Use of the MMPI-2 Content Scales, MMPI-2 Dep: Second version of the Minnesota Multiphasic Personality Inventory - Depression subscale in keeping with Hathaway et al: Minnesota Multiphasic Personality Inventory-2. Manual for Administration and Scoring, MOPSY: Monitoring trends and determinants in cardiovascular disease project psychosocial Interview Depression scale, NR: Not reported, OR: Odds ratio, PRIME-MD: Primary Care Evaluation of Mental Disorders, Py: Person years

Table A 7: Results of studies on the risk of myocardial infarction per one unit increase on depressive symptom rating scales

Author(s), year	Definition depression	Sample size (n / N)	Exposure group (n / N)	Comparison group (n / N)	Effect measure	Effect estimate* (unadjusted†)	Effect estimate* (most adjusted)
Ariyo et al (2000)	Depressive symptoms (10 item CES-D 10 per 5-unit increase in baseline score)	270 / 4,493	NA	NA	HR	1.06 (0.93 – 1.21)	1.12 (0.97 – 1.29)
	Depressive symptoms (10 item CES-D per 5-unit increase in mean cumulative score)					1.12 (0.96 – 1.29)	1.14 (0.98 – 1.34)
Barefoot & Schroll (1996)	Depressive symptoms (MMPI OBD subscale per 2 SD increase)	113 / 675	NA	NA	HR	1.71 (1.19 – 2.44)‡	1.70 (1.23 – 2.34)
Langvik & Hjemdal (2015)	Depressive symptoms (HADS-D per 1-unit increase)	770 / 28,859	NA	NA	OR		1.04 (1.01 – 1.04)
Pössel et al (2015)	Depressive symptoms (HPL per 1-point increase)	384 / 2,005	NA	NA	HR		1.04 (0.98 – 1.11)
Stürmer et al (2006)	Factor analysis of personality scales (per 1 SD increase)	72 / 32,875 py	NA	NA	HR	1.08 (0.86 – 1.35)	1.09 (0.84 – 1.4)

* Data are presented as estimate (95% confidence interval)
† Unless stated otherwise
‡ Age and sex adjusted
CES-D: Centre for Epidemiologic Studies Depression Scale, HADS-D: Hospital Anxiety and Depression Scale - Depression subscale, HPL: Human Population Laboratory Depression scale, HR: Hazard ratio, MI: Myocardial infarction, MMPI OBD subscale: Minnesota Multiphasic Personality Inventory - obvious depression subscale, NA: Not applicable, OR: Odds ratio, Py: Person years, SD: Standard deviation

Table A 8: Additional information on other covariates in included studies

Author	Other covariates
Arbelaez et al (2007)	Haemorrhagic stroke during follow-up†
Barefoot & Schroll (1996)	(Possible ischaemia†)
Bos et al (2008)	Intima-media thickness
Chi et al (2014)	Charlson comorbidity index
Cummings et al (2016)	Health insurance
Daskalopoulou et al (2016)	GP practice
Gilsanz et al (2015)	Baseline depressive symptoms††, participation††, sampling††, survival††
Gilsanz et al (2017)	Cumulative exposure to depression††, participation††, survival††
Glymour et al (2010)	Southern birth
Glymour et al (2012)	Southern birth
Henderson et al (2013)	Hip fractures
Joyce (2015)	K10***, other languages***, depression/anxiety treatment last month***, QOL***
Kim et al (2011)	Optimism, lung disease and arthritis†††, self-rated health
Kim et al (2013)	Purpose in life
Köhler et al (2013)	ApoE status
Kubzansky et al (2006)	Any known medical conditions†
Majed et al (2012)	Treatment for diabetes
Marijnissen et al (2014)	Neuroticism
Mejía-Lancheros et al (2014)	Intervention group, major endocrine or neurological disorders†
Moise et al Moise et al (2016)	Albumin to creatinine ratio, QT interval corrected for HR, health insurance
O'Brien et al (2015)	Coping skills, dialysis§§§
Pössel et al (2015)	Hopelessness
Scherrer et al (2010)	Hypothyroidism†
Scherrer et al (2011)	Health insurance
Sesso et al (1998)	Any known medical condition†
Sun et al (2016)	(Neurasthenia)

† Restriction

†† Variable was taken into account by applying inverse probability weights

*** Added in stepwise but did not end up in final model

††† Added as index of chronic illness (self-reported doctor diagnosis of high blood pressure, cancer, lung disease, psychiatric problems, and arthritis)

§§§ These variables were taken into account by fitting a separate model to obtain a stroke risk score which in turn was included in the regression model

GP: general practitioner; HR: heart rate; K10: Kessler Psychological Distress Scale; QOL: quality of life

A.II. UK Biobank

Table A 9: Results of the global cox.zph test of all fully adjusted models (complete cases)*

Outcome	Exposure	p-value
Major cardiovascular events	Depression	0.33
	Antidepressant use	0.33
	Hospital diagnosis with depression	0.31
	Self-reported depression	0.32
Myocardial infarction	Depression	0.45
	Antidepressant use	0.47
	Hospital diagnosis with depression	0.48
	Self-reported depression	0.49
Stroke	Depression	0.09
	Antidepressant use	0.10
	Hospital diagnosis with depression	0.07
	Self-reported depression	0.10
Major cardiovascular events	Depression and/ or hypertension	0.38
	Depression and/ or diabetes	0.37
	Depression and/ or high cholesterol levels	0.25
	Depression and/ or low educational attainment	0.38
	Depression and/ or high area-based deprivation	0.26

* The null hypothesis of the cox.zph function is that there are constant regression coefficients of the covariates over time. Therefore, a p-value <0.05 indicates a likely violation of the proportional hazard assumption

Table A 10: Results of the cox.zph test of the fully adjusted model investigating the association between depression and MCVE (complete cases)*

Variable	rho	chisq	p-value
Depression	-0.003	2.87E-02	0.87
Sex	0.001	1.75E-03	0.97
Age (2nd quartile)	-0.026	2.86E+00	0.09
Age (3rd quartile)	-0.013	7.69E-01	0.38
Age (4th quartile)	-0.031	4.14E+00	0.04
Ethnicity (Other ethnic groups)	-0.010	4.12E-01	0.52
Educational attainment (A levels, O levels, CSE, NVQ, or equivalent	-0.010	4.00E-01	0.53
Educational attainment (None of the above)	-0.025	2.68E+00	0.10
Area-based deprivation (2)	-0.017	1.28E+00	0.26
Area-based deprivation (3)	<0.001	1.84E-04	0.99
Area-based deprivation (4)	-0.020	1.60E+00	0.21
Area-based deprivation (5 - Most deprived)	0.005	8.75E-02	0.77
Income (52,000 to 100,000)	0.009	3.26E-01	0.57
Income (31,000 to 51,999)	0.001	3.24E-03	0.95
Income (18,000 to 30,999)	0.011	4.95E-01	0.48
Income (less than 18,000)	0.008	2.81E-01	0.60
Body mass index (overweight)	-0.016	1.05E+00	0.30
Body mass index (obese)	-0.025	2.63E+00	0.11
Body mass index (severely obese)	0.008	3.05E-01	0.58
Body mass index (morbidly obese)	-0.025	2.58E+00	0.11
Physical activity (moderate)	-0.023	2.15E+00	0.14
Physical activity (low)	-0.002	2.28E-02	0.88
Alcohol intake (risky drinking)	0.027	3.14E+00	0.08
Smoking status (previous)	-0.018	1.34E+00	0.25
Smoking status (current)	-0.021	1.82E+00	0.18
Fruit and vegetable intake (less than five a day)	-0.022	1.99E+00	0.16
Oily fish intake (less than once per week)	-0.001	1.03E-03	0.97
Oily fish intake (never)	-0.002	2.11E-02	0.88
Hypertension	0.007	2.19E-01	0.64
Diabetes	-0.002	1.57E-02	0.90
Cholesterol	0.010	3.96E-01	0.53
Family history of cardiovascular disease	0.012	6.19E-01	0.43
Family history of depression	0.006	1.34E-01	0.71
GLOBAL	NA	3.59E+01	0.33

CSE: Certificate of secondary education, NVQ: National vocational qualification

* The null hypothesis of the cox.zph function is that there are constant regression coefficients of the covariates over time. Therefore, a p-value <0.05 indicates a likely violation of the proportional hazard assumption

Table A 11: Convergence plots of multiple imputations in keeping with analyses on depression and MCVE

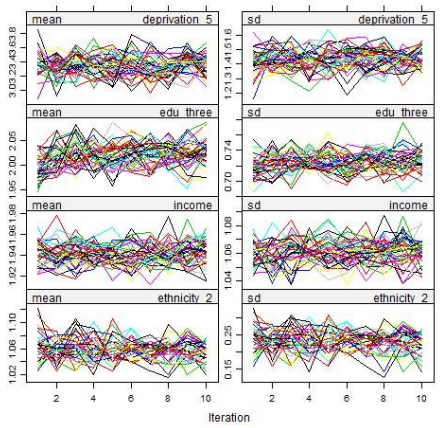
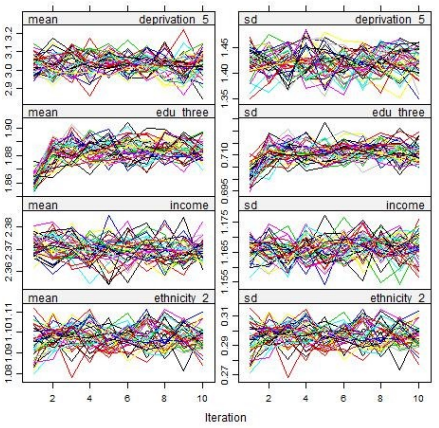
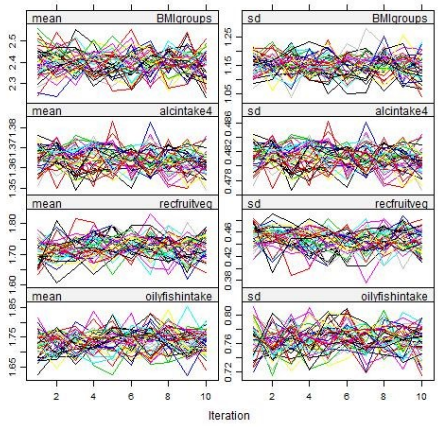
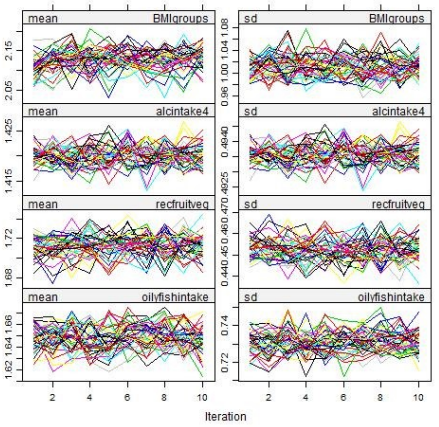
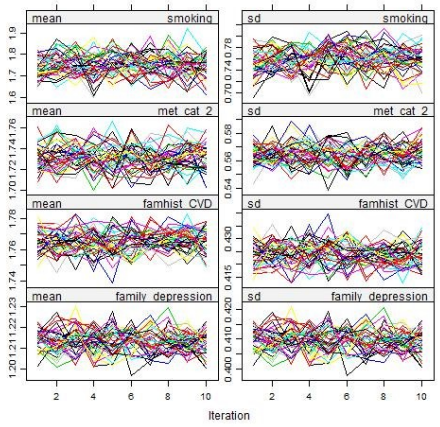
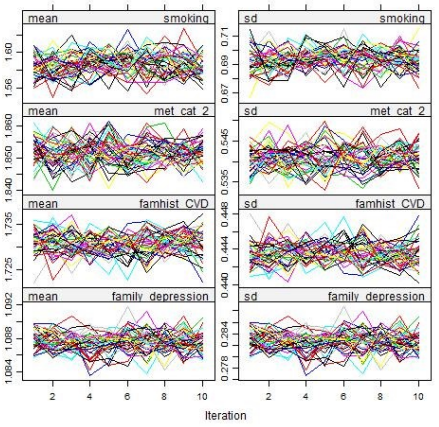
Exposure status: Depression Outcome: Major cardiovascular events Interaction: Comorbidities	**Exposure status: No depression** Outcome: Major cardiovascular events Interaction: Comorbidities
	
	
	

*Table A 12: Baseline characteristics separately for participants with and without antidepressant use at baseline (page 1 of 2)**

	No antidepressant use (n = 434,085, 93.3%)	**Antidepressant use** (n = 31,130, 6.7%)
Male (%)	196,653 (45.3)	8,725 (28.0)
Age (median [IQR])	57.0 [50.0, 63.0]	57.0 [50.0, 62.0]
Ethnicity (%)		
White	407,929 (94.0)	29,976 (96.3)
Other ethnic groups	23,768 (5.5)	1,019 (3.3)
Income (%)		
Greater than 100,000	21,557 (5.0)	700 (2.2)
52,000 to 100,000	79,895 (18.4)	3,367 (10.8)
31,000 to 51,999	99,510 (22.9)	5,830 (18.7)
18,000 to 30,999	93,026 (21.4)	6,936 (22.3)
Less than 18,000	75,030 (17.3)	9,285 (29.8)
Highest educational attainment (%)		
College or university degree	145,073 (33.4)	8,069 (25.9)
A levels, O levels, NVQ, or equivalent	213,253 (49.1)	15,800 (50.8)
None of the above	67,193 (15.5)	6,683 (21.5)
Area-based deprivation (%)		
1 (Least deprived)	89,533 (20.6)	5,418 (17.4)
2	88,102 (20.3)	5,626 (18.1)
3	87,657 (20.2)	5,956 (19.1)
4	86,450 (19.9)	6,388 (20.5)
5 (Most deprived)	81,825 (18.8)	7,685 (24.7)
Body mass index (%)		
Under- or normal weight	149,242 (34.4)	8,530 (27.4)
Overweight	184,207 (42.4)	12,165 (39.1)
Obese	71,647 (16.5)	6,502 (20.9)
Severely obese	19,403 (4.5)	2,463 (7.9)
Morbidly obese	7,134 (1.6)	1,261 (4.1)
Physical activity (%)		
High	40,313 (9.3)	2,014 (6.5)
Moderate	286,428 (66.0)	18,332 (58.9)
Low	89,944 (20.7)	9,223 (29.6)
Alcohol intake (%)		
Safe drinking	193,210 (44.5)	15,279 (49.1)
Risky drinking	178,299 (41.1)	9,610 (30.9)
Smoking status (%)		
Never	243,507 (56.1)	15,306 (49.2)
Previous	145,379 (33.5)	10,772 (34.6)
Current	42,761 (9.9)	4,891 (15.7)
Fruit and vegetable intake per day (%)		
At least five a day	130,907 (30.2)	9,731 (31.3)
Less than five a day	301,739 (69.5)	21,304 (68.4)

*Table A 12 continued: Baseline characteristics separately for participants with and without antidepressant use at baseline (page 2 of 2)**

	No antidepressant use (n = 434,085, 93.3%)	**Antidepressant use** (n = 31,130, 6.7%)
Oily fish intake (%)		
At least once a week	240,638 (55.4)	15,847 (50.9)
Less than once per week	143,957 (33.2)	10,548 (33.9)
Never	46,109 (10.6)	4,475 (14.4)
Hypertension (%)	235,034 (54.1)	17,629 (56.6)
Diabetes (%)	17,993 (4.1)	2,101 (6.7)
High cholesterol levels (%)	62,311 (14.4)	6,479 (20.8)
Family history of cardiovascular disease (%)		
No	103,170 (23.8)	6,414 (20.6)
Yes	297,832 (68.6)	22,302 (71.6)
Family history of depression (%)		
No	342,343 (78.9)	21,173 (68.0)
Yes	35,588 (8.2)	5,516 (17.7)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

*Table A 13: Baseline characteristics separately for participants with and without hospital diagnosis with depression at baseline (page 1 of 2)**

	No hospital diagnosis with depression (n = 461,984, 99.3%)	**Hospital diagnosis with depression** (n = 3,231, 0.7%)
Male (%)	204,190 (44.2)	1,188 (36.8)
Age (median [IQR])	57.0 [50.0, 63.0]	55.0 [48.0, 62.0]
Ethnicity (%)		
White	434,843 (94.1)	3,062 (94.8)
Other ethnic groups	24,641 (5.3)	146 (4.5)
Income (%)		
Greater than 100,000	22,225 (4.8)	32 (1.0)
52,000 to 100,000	83,040 (18.0)	222 (6.9)
31,000 to 51,999	104,885 (22.7)	455 (14.1)
18,000 to 30,999	99,305 (21.5)	657 (20.3)
Less than 18,000	83,036 (18.0)	1,279 (39.6)
Highest educational attainment (%)		
College or university degree	152,402 (33.0)	740 (22.9)
A levels, O levels, NVQ, or equivalent	227,416 (49.2)	1,637 (50.7)
None of the above	73,095 (15.8)	781 (24.2)
Area-based deprivation (%)		
1 (Least deprived)	94,558 (20.5)	393 (12.2)
2	93,268 (20.2)	460 (14.2)
3	93,104 (20.2)	509 (15.8)
4	92,154 (19.9)	684 (21.2)
5 (Most deprived)	88,332 (19.1)	1,178 (36.5)
Body mass index (%)		
Under- or normal weight	156,949 (34.0)	823 (25.5)
Overweight	195,180 (42.2)	1,192 (36.9)
Obese	77,441 (16.8)	708 (21.9)
Severely obese	21,581 (4.7)	285 (8.8)
Morbidly obese	8,225 (1.8)	170 (5.3)
Physical activity (%)		
High	42,127 (9.1)	200 (6.2)
Moderate	302,862 (65.6)	1,898 (58.7)
Low	98,217 (21.3)	950 (29.4)
Alcohol intake (%)		
Safe drinking	206,855 (44.8)	1,634 (50.6)
Risky drinking	186,922 (40.5)	987 (30.5)
Smoking status (%)		
Never	257,376 (55.7)	1,437 (44.5)
Previous	155,123 (33.6)	1,028 (31.8)
Current	46,912 (10.2)	740 (22.9)
Fruit and vegetable intake per day (%)		
At least five a day	139,714 (30.2)	924 (28.6)
Less than five a day	320,758 (69.4)	2,285 (70.7)

*Table A 13 continued: Baseline characteristics separately for participants with and without hospital diagnosis with depression at baseline (page 2 of 2)**

	No hospital diagnosis with depression (n = 461,984, 99.3%)	**Hospital diagnosis with depression** (n = 3,231, 0.7%)
Oily fish intake (%)		
At least once a week	254,907 (55.2)	1,578 (48.8)
Less than once per week	153,461 (33.2)	1,044 (32.3)
Never	50,024 (10.8)	560 (17.3)
Hypertension (%)	250,766 (54.3)	1,897 (58.7)
Diabetes (%)	19,828 (4.3)	266 (8.2)
High cholesterol levels (%)	68,107 (14.7)	683 (21.1)
Family history of cardiovascular disease (%)		
No	108,935 (23.6)	649 (20.1)
Yes	317,876 (68.8)	2,258 (69.9)
Family history of depression (%)		
No	361,498 (78.2)	2,018 (62.5)
Yes	40,464 (8.8)	640 (19.8)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified.

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

*Table A 14: Baseline characteristics separately for participants with and without self-reported depression at baseline (page 1 of 2)**

	No self-reported depression (n = 440,076, 94.6%)	**Self-reported depression** (n = 25,139, 5.4%)
Male (%)	197,281 (44.8)	8,097 (32.2)
Age (median [IQR])	57.0 [50.0, 63.0]	56.0 [49.0, 61.0]
Ethnicity (%)		
White	413,798 (94.0)	24,107 (95.9)
Other ethnic groups	23,888 (5.4)	899 (3.6)
Income (%)		
Greater than 100,000	21,613 (4.9)	644 (2.6)
52,000 to 100,000	80,305 (18.2)	2,957 (11.8)
31,000 to 51,999	100,291 (22.8)	5,049 (20.1)
18,000 to 30,999	94,329 (21.4)	5,633 (22.4)
Less than 18,000	76,993 (17.5)	7,322 (29.1)
Highest educational attainment (%)		
College or university degree	145,377 (33.0)	7,765 (30.9)
A levels, O levels, NVQ, or equivalent	216,457 (49.2)	12,596 (50.1)
None of the above	69,385 (15.8)	4,491 (17.9)
Area-based deprivation (%)		
1 (Least deprived)	90,766 (20.6)	4,185 (16.6)
2	89,240 (20.3)	4,488 (17.9)
3	88,915 (20.2)	4,698 (18.7)
4	87,609 (19.9)	5,229 (20.8)
5 (Most deprived)	83,016 (18.9)	6,494 (25.8)
Body mass index (%)		
Under- or normal weight	150,433 (34.2)	7,339 (29.2)
Overweight	186,455 (42.4)	9,917 (39.4)
Obese	73,135 (16.6)	5,014 (19.9)
Severely obese	20,055 (4.6)	1,811 (7.2)
Morbidly obese	7,467 (1.7)	928 (3.7)
Physical activity (%)		
High	40,535 (9.2)	1,792 (7.1)
Moderate	289,435 (65.8)	15,325 (61.0)
Low	92,276 (21.0)	6,891 (27.4)
Alcohol intake (%)		
Safe drinking	196,730 (44.7)	11,759 (46.8)
Risky drinking	179,285 (40.7)	8,624 (34.3)
Smoking status (%)		
Never	246,398 (56.0)	12,415 (49.4)
Previous	147,647 (33.6)	8,504 (33.8)
Current	43,543 (9.9)	4,109 (16.3)
Fruit and vegetable intake per day (%)		
At least five a day	133,414 (30.3)	7,224 (28.7)
Less than five a day	305,203 (69.4)	17,840 (71.0)

*Table A 14 continued: Baseline characteristics separately for participants with and without self-reported depression at baseline (page 2 of 2)**

	No self-reported depression (n = 440,076, 94.6%)	**Self-reported depression** (n = 25,139, 5.4%)
Oily fish intake (%)		
At least once a week	243,841 (55.4)	12,644 (50.3)
Less than once per week	145,836 (33.1)	8,669 (34.5)
Never	46,962 (10.7)	3,622 (14.4)
Hypertension (%)	239,578 (54.4)	13,085 (52.1)
Diabetes (%)	18,804 (4.3)	1,290 (5.1)
High cholesterol levels (%)	64,549 (14.7)	4,241 (16.9)
Family history of cardiovascular disease (%)		
No	104,359 (23.7)	5,225 (20.8)
Yes	302,198 (68.7)	17,936 (71.3)
Family history of depression (%)		
No	347,293 (78.9)	16,223 (64.5)
Yes	35,754 (8.1)	5,350 (21.3)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified.

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

*Table A 15: Baseline characteristics separately for participants with and without MCVE during follow-up (page 1 of 2)**

	No major cardiovascular event (n = 457,366, 98.3%)	**Major cardiovascular event** (n = 7,849, 1.7%)
Male (%)	200,250 (43.8)	5,128 (65.3)
Age (median [IQR])	57.0 [50.0, 63.0]	62.0 [56.0, 66.0]
Ethnicity (%)		
White	430,474 (94.1)	7,431 (94.7)
Other ethnic groups	24,430 (5.3)	357 (4.5)
Income (%)		
Greater than 100,000	22,071 (4.8)	186 (2.4)
52,000 to 100,000	82,408 (18.0)	854 (10.9)
31,000 to 51,999	103,853 (22.7)	1,487 (18.9)
18,000 to 30,999	98,089 (21.4)	1,873 (23.9)
Less than 18,000	82,236 (18.0)	2,079 (26.5)
Highest educational attainment (%)		
College or university degree	151,255 (33.1)	1,887 (24.0)
A levels, O levels, NVQ, or equivalent	225,405 (49.3)	3,648 (46.5)
None of the above	71,814 (15.7)	2,062 (26.3)
Area-based deprivation (%)		
1 (Least deprived)	93,488 (20.4)	1,463 (18.6)
2	92,270 (20.2)	1,458 (18.6)
3	92,020 (20.1)	1,593 (20.3)
4	91,326 (20.0)	1,512 (19.3)
5 (Most deprived)	87,695 (19.2)	1,815 (23.1)
Body mass index (%)		
Under- or normal weight	155,833 (34.1)	1,939 (24.7)
Overweight	192,883 (42.2)	3,489 (44.5)
Obese	76,470 (16.7)	1,679 (21.4)
Severely obese	21,407 (4.7)	459 (5.8)
Morbidly obese	8,198 (1.8)	197 (2.5)
Physical activity (%)		
High	41,785 (9.1)	542 (6.9)
Moderate	299,741 (65.5)	5,019 (63.9)
Low	97,292 (21.3)	1,875 (23.9)
Alcohol intake (%)		
Safe drinking	205,285 (44.9)	3,204 (40.8)
Risky drinking	184,607 (40.4)	3,302 (42.1)
Smoking status (%)		
Never	255,480 (55.9)	3,333 (42.5)
Previous	153,229 (33.5)	2,922 (37.2)
Current	46,120 (10.1)	1,532 (19.5)
Fruit and vegetable intake per day (%)		
At least five a day	138,525 (30.3)	2,113 (26.9)
Less than five a day	317,344 (69.4)	5,699 (72.6)

*Table A 15 continued: Baseline characteristics separately for participants with and without MCVE during follow-up (page 2 of 2)**

	No major cardiovascular event (n = 457,366, 98.3%)	**Major cardiovascular event** (n = 7,849, 1.7%)
Oily fish intake (%)		
At least once a week	252,271 (55.2)	4,214 (53.7)
Less than once per week	151,955 (33.2)	2,550 (32.5)
Never	49,590 (10.8)	994 (12.7)
Hypertension (%)	246,658 (53.9)	6,005 (76.5)
Diabetes (%)	19,326 (4.2)	768 (9.8)
High cholesterol levels (%)	66,853 (14.6)	1,937 (24.7)
Family history of cardiovascular disease (%)		
No	107,917 (23.6)	1,667 (21.2)
Yes	314,722 (68.8)	5,412 (69.0)
Family history of depression (%)		
No	357,737 (78.2)	5,779 (73.6)
Yes	40,471 (8.8)	633 (8.1)
Depression (%)	39530 (8.6)	823 (10.5)
Antidepressant use (%)	30475 (6.7)	655 (8.3)
Self-reported depression (%)	24666 (5.4)	473 (6.0)
Hospital diagnosis with depression (%)	3136 (0.7)	95 (1.2)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified.

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

*Table A 16: Baseline characteristics separately for participants with and without stroke during follow-up (page 1 of 2)**

	No stroke (n = 462,090, 99.3%)	**Stroke** (n = 3,125, 0.7%)
Male (%)	203,624 (44.1)	1,754 (56.1)
Age (median [IQR])	57.0 [50.0, 63.0]	63.0 [57.0, 66.0]
Ethnicity (%)		
White	434,933 (94.1)	2,972 (95.1)
Other ethnic groups	24,652 (5.3)	135 (4.3)
Income (%)		
Greater than 100,000	22,185 (4.8)	72 (2.3)
52,000 to 100,000	82,956 (18.0)	306 (9.8)
31,000 to 51,999	104,787 (22.7)	553 (17.7)
18,000 to 30,999	99,216 (21.5)	746 (23.9)
Less than 18,000	83,441 (18.1)	874 (28.0)
Highest educational attainment (%)		
College or university degree	152,363 (33.0)	779 (24.9)
A levels, O levels, NVQ, or equivalent	227,617 (49.3)	1,436 (46.0)
None of the above	73,061 (15.8)	815 (26.1)
Area-based deprivation (%)		
1 (Least deprived)	94,398 (20.4)	553 (17.7)
2	93,162 (20.2)	566 (18.1)
3	92,982 (20.1)	631 (20.2)
4	92,221 (20.0)	617 (19.7)
5 (Most deprived)	88,759 (19.2)	751 (24.0)
Body mass index (%)		
Under- or normal weight	156,907 (34.0)	865 (27.7)
Overweight	195,045 (42.2)	1,327 (42.5)
Obese	77,529 (16.8)	620 (19.8)
Severely obese	21,681 (4.7)	185 (5.9)
Morbidly obese	8,307 (1.8)	88 (2.8)
Physical activity (%)		
High	42,099 (9.1)	228 (7.3)
Moderate	302,739 (65.5)	2,021 (64.7)
Low	98,447 (21.3)	720 (23.0)
Alcohol intake (%)		
Safe drinking	207,207 (44.8)	1,282 (41.0)
Risky drinking	186,580 (40.4)	1,329 (42.5)
Smoking status (%)		
Never	257,387 (55.7)	1,426 (45.6)
Previous	155,014 (33.5)	1,137 (36.4)
Current	47,114 (10.2)	538 (17.2)
Fruit and vegetable intake per day (%)		
At least five a day	139,763 (30.2)	875 (28.0)
Less than five a day	320,804 (69.4)	2,239 (71.6)

*Table A 16 continued: Baseline characteristics separately for participants with and without stroke during follow-up (page 2 of 2)**

	No stroke (n = 462,090, 99.3%)	**Stroke** (n = 3,125, 0.7%)
Oily fish intake (%)		
At least once a week	254,760 (55.1)	1,725 (55.2)
Less than once per week	153,493 (33.2)	1,012 (32.4)
Never	50,222 (10.9)	362 (11.6)
Hypertension (%)	250,293 (54.2)	2,370 (75.8)
Diabetes (%)	19,776 (4.3)	318 (10.2)
High cholesterol levels (%)	68,015 (14.7)	775 (24.8)
Family history of cardiovascular disease (%)		
No	108,906 (23.6)	678 (21.7)
Yes	318,018 (68.8)	2,116 (67.7)
Family history of depression (%)		
No	361,186 (78.2)	2,330 (74.6)
Yes	40,860 (8.8)	244 (7.8)
Depression (%)	40,025 (8.7)	328 (10.5)
Antidepressant use (%)	30,858 (6.7)	272 (8.7)
Self-reported depression (%)	24,951 (5.4)	188 (6.0)
Hospital diagnosis with depression (%)	3,199 (0.7)	32 (1.0)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified.

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

*Table A 17: Baseline characteristics separately for participants with and without MI during follow-up (page 1 of 2)**

	No myocardial infarction (n = 460,330, 98.9%)	**Myocardial infarction** (n = 4,885, 1.1%)
Male (%)	201,888 (43.9)	3,490 (71.4)
Age (median [IQR])	57.0 [50.0, 63.0]	61.0 [56.0, 65.0]
Ethnicity (%)		
White	433,291 (94.1)	4,614 (94.5)
Other ethnic groups	24,559 (5.3)	228 (4.7)
Income (%)		
Greater than 100,000	22,142 (4.8)	115 (2.4)
52,000 to 100,000	82,703 (18.0)	559 (11.4)
31,000 to 51,999	104,380 (22.7)	960 (19.7)
18,000 to 30,999	98,795 (21.5)	1,167 (23.9)
Less than 18,000	83,057 (18.0)	1,258 (25.8)
Highest educational attainment (%)		
College or university degree	151,998 (33.0)	1,144 (23.4)
A levels, O levels, NVQ, or equivalent	226,765 (49.3)	2,288 (46.8)
None of the above	72,582 (15.8)	1,294 (26.5)
Area-based deprivation (%)		
1 (Least deprived)	94,017 (20.4)	934 (19.1)
2	92,807 (20.2)	921 (18.9)
3	92,615 (20.1)	998 (20.4)
4	91,910 (20.0)	928 (19.0)
5 (Most deprived)	88,408 (19.2)	1,102 (22.6)
Body mass index (%)		
Under- or normal weight	156,658 (34.0)	1,114 (22.8)
Overweight	194,144 (42.2)	2,228 (45.6)
Obese	77,056 (16.7)	1,093 (22.4)
Severely obese	21,582 (4.7)	284 (5.8)
Morbidly obese	8,278 (1.8)	117 (2.4)
Physical activity (%)		
High	42,005 (9.1)	322 (6.6)
Moderate	301,659 (65.5)	3,101 (63.5)
Low	97,969 (21.3)	1,198 (24.5)
Alcohol intake (%)		
Safe drinking	206,501 (44.9)	1,988 (40.7)
Risky drinking	185,863 (40.4)	2,046 (41.9)
Smoking status (%)		
Never	256,838 (55.8)	1,975 (40.4)
Previous	154,303 (33.5)	1,848 (37.8)
Current	46,631 (10.1)	1,021 (20.9)
Fruit and vegetable intake per day (%)		
At least five a day	139,354 (30.3)	1,284 (26.3)
Less than five a day	319,468 (69.4)	3,575 (73.2)

*Table A 17 continued: Baseline characteristics separately for participants with and without MI during follow-up (page 2 of 2)**

	No myocardial infarction (n = 460,330, 98.9%)	**Myocardial infarction** (n = 4,885, 1.1%)
Oily fish intake (%)		
At least once a week	253,914 (55.2)	2,571 (52.6)
Less than once per week	152,915 (33.2)	1,590 (32.5)
Never	49,928 (10.8)	656 (13.4)
Hypertension (%)	248,889 (54.1)	3,774 (77.3)
Diabetes (%)	19,611 (4.3)	483 (9.9)
High cholesterol levels (%)	67,570 (14.7)	1,220 (25.0)
Family history of cardiovascular disease (%)		
No	108,560 (23.6)	1,024 (21.0)
Yes	316,731 (68.8)	3,403 (69.7)
Family history of depression (%)		
No	359,949 (78.2)	3,567 (73.0)
Yes	40,704 (8.8)	400 (8.2)
Depression (%)	39,838 (8.7)	515 (10.5)
Antidepressant use (%)	30,731 (6.7)	399 (8.2)
Self-reported depression (%)	24,841 (5.4)	298 (6.1)
Hospital diagnosis with depression (%)	3,166 (0.7)	65 (1.3)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism.
Data are given as n (%) unless specified.
CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

Table A 18: Hazard ratios (95% CI) of the association between depression and MCVE (imputed data)

	Unadjusted HR (95% CI)	**Adjusted HR (95% CI) Model 1**	**Adjusted HR (95% CI) Model 2**
Depression	1.24 (1.15 – 1.33)	1.35 (1.25 – 1.45)	1.21 (1.12 – 1.30)
Antidepressant use	1.27 (1.18 – 1.38)	1.38 (1.28 – 1.50)	1.22 (1.12 – 1.32)
Hospital diagnosis with depression	1.81 (1.48 – 2.22)	1.78 (1.45 – 2.18)	1.52 (1.24 – 1.86)
Self-reported depression	1.12 (1.02 – 1.23)	1.27 (1.16 – 1.40)	1.16 (1.05 – 1.27)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

HR: Hazard ratio

Table A 19: Hazard ratios (95% CI) of the association between depression and stroke and MI (imputed data)

	Unadjusted HR (95% CI)	**Adjusted HR (95% CI) Model 1**	**Adjusted HR (95% CI) Model 2**
Stroke	1.24 (1.10 – 1.39)	1.28 (1.14 – 1.43)	1.18 (1.05 – 1.33)
Myocardial infarction	1.24 (1.13 – 1.36)	1.41 (1.29 – 1.55)	1.24 (1.12 – 1.36)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

HR: Hazard ratio

*Table A 20: Baseline characteristics for participants with depression and/ or hypertension at baseline (page 1 of 2)**

	No depression, no hypertension (n=194,547, 41.8%)	**Depression, no hypertension** (n=18,005, 3.9%)	**No depression, hypertension** (n=230,315, 49.5%)	**Depression, hypertension** (n=22,348, 4.8%)
Male (%)	75,002 (38.6)	4,399 (24.4)	118,119 (51.3)	7,858 (35.2)
Age (median [IQR])	53.0 [47.0, 60.0]	53.0 [47.0, 60.0]	60.0 [53.0, 64.0]	59.0 [52.0, 64.0]
Ethnicity (%)				
White	181,945 (93.5)	17,275 (95.9)	217,193 (94.3)	21,492 (96.2)
Other ethnic groups	11,358 (5.8)	644 (3.6)	12,040 (5.2)	745 (3.3)
Income (%)				
Greater than 100,000	12,767 (6.6)	573 (3.2)	8,541 (3.7)	376 (1.7)
52,000 to 100,000	43,169 (22.2)	2,521 (14.0)	35,522 (15.4)	2,050 (9.2)
31,000 to 51,999	47,555 (24.4)	3,902 (21.7)	50,038 (21.7)	3,845 (17.2)
18,000 to 30,999	37,737 (19.4)	3,887 (21.6)	53,246 (23.1)	5,092 (22.8)
Less than 18,000	26,931 (13.8)	4,672 (25.9)	45,591 (19.8)	7,121 (31.9)
Highest educational attainment (%)				
College or university degree	75,325 (38.7)	5,816 (32.3)	66,532 (28.9)	5,469 (24.5)
A levels, O levels, CSE, NVQ, or equivalent	94,716 (48.7)	9,227 (51.2)	114,138 (49.6)	10,972 (49.1)
None of the above	20,625 (10.6)	2,691 (14.9)	45,107 (19.6)	5,453 (24.4)
Area-based deprivation (%)				
1 (Least deprived)	40,340 (20.7)	3,199 (17.8)	47,659 (20.7)	3,753 (16.8)
2	39,051 (20.1)	3,279 (18.2)	47,415 (20.6)	3,983 (17.8)
3	38,996 (20.0)	3,404 (18.9)	46,987 (20.4)	4,226 (18.9)
4	39,316 (20.2)	3,692 (20.5)	45,233 (19.6)	4,597 (20.6)
5 (Most deprived)	36,607 (18.8)	4,398 (24.4)	42,755 (18.6)	5,750 (25.7)
Body mass index (%)				
Under- or normal weight	88,846 (45.7)	7,076 (39.3)	57,355 (24.9)	4,495 (20.1)
Overweight	77,379 (39.8)	7,127 (39.6)	103,141 (44.8)	8,725 (39.0)
Obese	21,435 (11.0)	2,670 (14.8)	48,533 (21.1)	5,511 (24.7)
Severely obese	4,453 (2.3)	757 (4.2)	14,429 (6.3)	2,227 (10.0)
Morbidly obese	1,152 (0.6)	275 (1.5)	5,756 (2.5)	1,212 (5.4)
Physical activity (%)				
High	19,933 (10.2)	1,333 (7.4)	19,663 (8.5)	1,398 (6.3)
Moderate	128,772 (66.2)	11,116 (61.7)	151,797 (65.9)	13,075 (58.5)
Low	38,799 (19.9)	4,816 (26.7)	48,925 (21.2)	6,627 (29.7)
Alcohol intake (%)				
Safe drinking	92,211 (47.4)	9,006 (50.0)	96,745 (42.0)	10,527 (47.1)
Risky drinking	73,307 (37.7)	5,474 (30.4)	101,640 (44.1)	7,488 (33.5)
Smoking status (%)				
Never	113,686 (58.4)	9,035 (50.2)	125,198 (54.4)	10,894 (48.7)
Previous	58,847 (30.2)	5,758 (32.0)	83,411 (36.2)	8,135 (36.4)
Current	20,842 (10.7)	3,132 (17.4)	20,495 (8.9)	3,183 (14.2)

*Table A 20 continued: Baseline characteristics for participants with depression and/ or hypertension at baseline (page 2 of 2)**

	No depression, no hypertension (n=194,547, 41.8%)	**Depression, no hypertension** (n=18,005, 3.9%)	**No depression, hypertension** (n=230,315, 49.5%)	**Depression, hypertension** (n=22,348, 4.8%)
Fruit and vegetable intake per day (%)				
At least five a day	58,203 (29.9)	5,383 (29.9)	70,082 (30.4)	6,970 (31.2)
Less than five a day	135,520 (69.7)	12,566 (69.8)	159,660 (69.3)	15,297 (68.4)
Oily fish intake (%)				
At least once a week	103,772 (53.3)	8,933 (49.6)	132,002 (57.3)	11,778 (52.7)
Less than once per week	67,526 (34.7)	6,336 (35.2)	73,345 (31.8)	7,298 (32.7)
Never	21,616 (11.1)	2,596 (14.4)	23,311 (10.1)	3,061 (13.7)
Diabetes (%)	3,098 (1.6)	418 (2.3)	14,511 (6.3)	2,067 (9.2)
High cholesterol levels (%)	11,671 (6.0)	1,582 (8.8)	49,384 (21.4)	6,153 (27.5)
Family history of cardiovascular disease (%)				
No	54,369 (27.9)	4,414 (24.5)	46,825 (20.3)	3,976 (17.8)
Yes	126,326 (64.9)	12,237 (68.0)	165,062 (71.7)	16,509 (73.9)
Family history of depression (%)				
No	155,584 (80.0)	12,238 (68.0)	180,827 (78.5)	14,867 (66.5)
Yes	16,860 (8.7)	3,488 (19.4)	16,849 (7.3)	3,907 (17.5)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified.

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

*Table A 21: Baseline characteristics for participants with depression and/ or diabetes (page 1 of 2)**

	No depression, no diabetes (n=407,253, 87.5%)	**Depression, no diabetes** (n=37,868, 8.1%)	**No depression, diabetes** (n=17,609, 3.9%)	**Depression, diabetes** (n=2,485, 0.5%)
Male (%)	182,209 (44.7)	11,106 (29.3)	10,912 (62.0)	1,151 (46.3)
Age (median [IQR])	57.0 [49.0, 63.0]	56.0 [49.0, 62.0]	61.0 [55.0, 65.0]	59.0 [53.0, 64.0]
Ethnicity (%)				
White	384,078 (94.3)	36,480 (96.3)	15,060 (85.5)	2,287 (92.0)
Other ethnic groups	21,002 (5.2)	1,208 (3.2)	2,396 (13.6)	181 (7.3)
Income (%)				
Greater than 100,000	20,868 (5.1)	925 (2.4)	440 (2.5)	24 (1.0)
52,000 to 100,000	76,672 (18.8)	4,408 (11.6)	2,019 (11.5)	163 (6.6)
31,000 to 51,999	94,433 (23.2)	7,405 (19.6)	3,160 (17.9)	342 (13.8)
18,000 to 30,999	86,867 (21.3)	8,437 (22.3)	4,116 (23.4)	542 (21.8)
Less than 18,000	67,885 (16.7)	10,824 (28.6)	4,637 (26.3)	969 (39.0)
Highest educational attainment (%)				
College or university degree	137,367 (33.7)	10,750 (28.4)	4,490 (25.5)	535 (21.5)
A levels, O levels, NVQ, or equivalent	200,545 (49.2)	19,024 (50.2)	8,309 (47.2)	1,175 (47.3)
None of the above	61,434 (15.1)	7,444 (19.7)	4,298 (24.4)	700 (28.2)
Area-based deprivation (%)				
1 (Least deprived)	85,280 (20.9)	6,637 (17.5)	2,719 (15.4)	315 (12.7)
2	83,501 (20.5)	6,886 (18.2)	2,965 (16.8)	376 (15.1)
3	82,677 (20.3)	7,240 (19.1)	3,306 (18.8)	390 (15.7)
4	80,832 (19.8)	7,758 (20.5)	3,717 (21.1)	531 (21.4)
5 (Most deprived)	74,488 (18.3)	9,278 (24.5)	4,874 (27.7)	870 (35.0)
Body mass index (%)				
Under- or normal weight	143,978 (35.4)	11,338 (29.9)	2,223 (12.6)	233 (9.4)
Overweight	174,288 (42.8)	15,178 (40.1)	6,232 (35.4)	674 (27.1)
Obese	64,723 (15.9)	7,447 (19.7)	5,245 (29.8)	734 (29.5)
Severely obese	16,507 (4.1)	2,523 (6.7)	2,375 (13.5)	461 (18.6)
Morbidly obese	5,556 (1.4)	1,146 (3.0)	1,352 (7.7)	341 (13.7)
Physical activity (%)				
High	38,301 (9.4)	2,612 (6.9)	1,295 (7.4)	119 (4.8)
Moderate	270,033 (66.3)	22,912 (60.5)	10,536 (59.8)	1,279 (51.5)
Low	82,874 (20.3)	10,520 (27.8)	4,850 (27.5)	923 (37.1)
Alcohol intake (%)				
Safe drinking	180,454 (44.3)	18,246 (48.2)	8,502 (48.3)	1,287 (51.8)
Risky drinking	169,090 (41.5)	12,389 (32.7)	5,857 (33.3)	573 (23.1)
Smoking status (%)				
Never	230,323 (56.6)	18,850 (49.8)	8,561 (48.6)	1,079 (43.4)
Previous	135,135 (33.2)	12,870 (34.0)	7,123 (40.5)	1,023 (41.2)
Current	39,593 (9.7)	5,950 (15.7)	1,744 (9.9)	365 (14.7)

*Table A 21 continued: Baseline characteristics for participants with depression and/ or diabetes (page 2 of 2)**

	No depression, no diabetes (n=407,253, 87.5%)	**Depression, no diabetes** (n=37,868, 8.1%)	**No depression, diabetes** (n=17,609, 3.9%)	**Depression, diabetes** (n=2,485, 0.5%)
Fruit and vegetable intake per day (%)				
At least five a day	122,116 (30.0)	11,481 (30.3)	6,169 (35.0)	872 (35.1)
Less than five a day	283,847 (69.7)	26,263 (69.4)	11,333 (64.4)	1,600 (64.4)
Oily fish intake (%)				
At least once a week	225,934 (55.5)	19,442 (51.3)	9,840 (55.9)	1,269 (51.1)
Less than once per week	135,507 (33.3)	12,821 (33.9)	5,364 (30.5)	813 (32.7)
Never	42,746 (10.5)	5,287 (14.0)	2,181 (12.4)	370 (14.9)
Hypertension (%)	215,804 (53.0)	20,281 (53.6)	14,511 (82.4)	2,067 (83.2)
High cholesterol levels (%)	47,758 (11.7)	5,762 (15.2)	13,297 (75.5)	1,973 (79.4)
Family history of cardiovascular disease (%)				
No	97,524 (23.9)	7,921 (20.9)	3,670 (20.8)	469 (18.9)
Yes	279,346 (68.6)	26,970 (71.2)	12,042 (68.4)	1,776 (71.5)
Family history of depression (%)				
No	323,354 (79.4)	25,453 (67.2)	13,057 (74.1)	1,652 (66.5)
Yes	32,583 (8.0)	7,022 (18.5)	1,126 (6.4)	373 (15.0)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

*Table A 22: Baseline characteristics for participants with depression and/ or high cholesterol levels (page 1 of 2)**

	No depression, low cholesterol (n=363,807, 78.2%)	**Depression, low cholesterol** (n=32,618, 7.0%)	**No depression, high cholesterol** (n=61,055, 13.1%)	**Depression, high cholesterol** (n=7,735, 1.7%)
Male (%)	158,012 (43.4)	9,099 (27.9)	35,109 (57.5)	3,158 (40.8)
Age (median [IQR])	56.0 [49.0, 62.0]	55.0 [49.0, 61.0]	62.0 [57.0, 66.0]	61.0 [56.0, 65.0]
Ethnicity (%)				
White	342,657 (94.2)	31,414 (96.3)	56,481 (92.5)	7,353 (95.1)
Other ethnic groups	19,152 (5.3)	1,057 (3.2)	4,246 (7.0)	0,332 (4.3)
Income (%)				
Greater than 100,000	19,162 (5.3)	840 (2.6)	2,146 (3.5)	109 (1.4)
52,000 to 100,000	71,161 (19.6)	4,023 (12.3)	7,530 (12.3)	548 (7.1)
31,000 to 51,999	86,039 (23.6)	6,621 (20.3)	11,554 (18.9)	1,126 (14.6)
18,000 to 30,999	76,465 (21.0)	7,266 (22.3)	14,518 (23.8)	1,713 (22.1)
Less than 18,000	57,696 (15.9)	8,968 (27.5)	14,826 (24.3)	2,825 (36.5)
Highest educational attainment (%)				
College or university degree	125,810 (34.6)	9,580 (29.4)	16,047 (26.3)	1,705 (22.0)
A levels, O levels, NVQ, or equivalent	180,014 (49.5)	16,639 (51.0)	28,840 (47.2)	3,560 (46.0)
None of the above	50,921 (14.0)	5,850 (17.9)	14,811 (24.3)	2,294 (29.7)
Area-based deprivation (%)				
1 (Least deprived)	76,428 (21.0)	5,771 (17.7)	11,571 (19.0)	1,181 (15.3)
2	74,558 (20.5)	5,998 (18.4)	11,908 (19.5)	1,264 (16.3)
3	73,818 (20.3)	6,217 (19.1)	12,165 (19.9)	1,413 (18.3)
4	72,321 (19.9)	6,700 (20.5)	12,228 (20.0)	1,589 (20.5)
5 (Most deprived)	66,254 (18.2)	7,866 (24.1)	13,108 (21.5)	2,282 (29.5)
Body mass index (%)				
Under- or normal weight	134,544 (37.0)	10,395 (31.9)	11,657 (19.1)	1,176 (15.2)
Overweight	153,068 (42.1)	12,894 (39.5)	27,452 (45.0)	2,958 (38.2)
Obese	55,010 (15.1)	6,066 (18.6)	14,958 (24.5)	2,115 (27.3)
Severely obese	14,231 (3.9)	2,056 (6.3)	4,651 (7.6)	928 (12.0)
Morbidly obese	4,943 (1.4)	999 (3.1)	1,965 (3.2)	488 (6.3)
Physical activity (%)				
High	34,826 (9.6)	2,304 (7.1)	4,770 (7.8)	427 (5.5)
Moderate	241,287 (66.3)	19,854 (60.9)	39,282 (64.3)	4,337 (56.1)
Low	73,375 (20.2)	8,939 (27.4)	14,349 (23.5)	2,504 (32.4)
Alcohol intake (%)				
Safe drinking	162,633 (44.7)	15,728 (48.2)	26,323 (43.1)	3,805 (49.2)
Risky drinking	148,878 (40.9)	10,630 (32.6)	26,069 (42.7)	2,332 (30.1)
Smoking status (%)				
Never	209,162 (57.5)	16,565 (50.8)	29,722 (48.7)	3,364 (43.5)
Previous	117,318 (32.2)	10,820 (33.2)	24,940 (40.8)	3,073 (39.7)
Current	35,333 (9.7)	5,075 (15.6)	6,004 (9.8)	1,240 (16.0)

*Table A 22 continued: Baseline characteristics for participants with depression and/ or high cholesterol levels (page 2 of 2)**

	No depression, low cholesterol (n=363,807, 78.2%)	**Depression, low cholesterol** (n=32,618, 7.0%)	**No depression, high cholesterol** (n=61,055, 13.1%)	**Depression, high cholesterol** (n=7,735, 1.7%)
Fruit and vegetable intake per day (%)				
At least five a day	108,653 (29.9)	9,769 (29.9)	19,632 (32.2)	2,584 (33.4)
Less than five a day	253,945 (69.8)	22,754 (69.8)	41,235 (67.5)	5,109 (66.1)
Oily fish intake (%)				
At least once a week	198,717 (54.6)	16,459 (50.5)	37,057 (60.7)	4,252 (55.0)
Less than once per week	122,914 (33.8)	11,251 (34.5)	17,957 (29.4)	2,383 (30.8)
Never	39,379 (10.8)	4,645 (14.2)	5,548 (9.1)	1,012 (13.1)
Hypertension (%)	180,931 (49.7)	16,195 (49.7)	49,384 (80.9)	6,153 (79.5)
Diabetes (%)	4,312 (1.2)	512 (1.6)	13,297 (21.8)	1,973 (25.5)
Family history of cardiovascular disease (%)				
No	90,476 (24.9)	7,096 (21.8)	10,718 (17.6)	1,294 (16.7)
Yes	246,201 (67.7)	22,965 (70.4)	45,187 (74.0)	5,781 (74.7)
Family history of depression (%)				
No	289,392 (79.5)	21,947 (67.3)	47,019 (77.0)	5,158 (66.7)
Yes	29,426 (8.1)	6,170 (18.9)	4,283 (7.0)	1,225 (15.8)

* This table does not show the number and percentage with missing values. For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

Table A 23: Hazard ratios (95% CI) of the association between depression and/ or different comorbidities and MCVE (imputed data)

	Unadjusted HR (95% CI)	**Adjusted HR (95% CI) Model 1**	**Adjusted HR (95% CI) Model 2**
No depression, no hypertension	ref.	ref.	ref.
Depression alone	1.12 (0.96 – 1.31)	1.17 (1.00 – 1.37)	1.06 (0.91 – 1.24)
Hypertension alone	2.74 (2.59 – 2.89)	1.84 (1.74 – 1.94)	1.78 (1.68 – 1.89)
Depression and hypertension	3.43 (3.14 – 3.76)	2.55 (2.32 – 2.79)	2.23 (2.03 – 2.45)
No diabetes, no depression	ref.	ref.	ref.
Depression alone	1.22 (1.13 – 1.32)	1.34 (1.24 – 1.45)	1.22 (1.13 – 1.32)
Diabetes alone	2.50 (2.31 – 2.71)	1.69 (1.55 – 1.83)	1.47 (1.35 – 1.61)
Depression and diabetes	2.68 (2.20 – 3.26)	2.05 (1.69 – 2.50)	1.63 (1.33 – 2.00)
No depression, low cholesterol levels	ref.	ref.	ref.
Depression alone	1.16 (1.06 – 1.26)	1.28 (1.18 – 1.40)	1.17 (1.07 – 1.28)
High cholesterol levels alone	1.90 (1.80 – 2.01)	1.22 (1.15 – 1.29)	0.97 (0.91 – 1.03)
Depression and high cholesterol levels	2.39 (2.12 – 2.71)	1.74 (1.54 – 1.97)	1.24 (1.09 – 1.41)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval, HR: Hazard ratio, ref.: reference

*Table A 24: Baseline characteristics for participants with depression and/ or low educational attainment (page 1 of 2)**

	No depression, high education (n=141,857, 31.1%)	**Depression, high education** (n=11,285, 2.5%)	**No depression, low education** (n=274,586, 60.2%)	**Depression, low education** (n=28,343, 6.2%)
Male (%)	67,480 (47.6)	3,802 (33.7)	121,717 (44.3)	8,236 (29.1)
Age (median [IQR])	55.0 [48.0, 61.0]	55.0 [49.0, 61.0]	58.0 [51.0, 64.0]	57.0 [50.0, 63.0]
Ethnicity (%)				
White	132,016 (93.1)	10,788 (95.6)	260,715 (94.9)	27,380 (96.6)
Other ethnic groups	9,250 (6.5)	451 (4.0)	13,081 (4.8)	881 (3.1)
Income (%)				
Greater than 100,000	15,385 (10.8)	664 (5.9)	5,912 (2.2)	283 (1.0)
52,000 to 100,000	43,110 (30.4)	2,476 (21.9)	35,493 (12.9)	2,089 (7.4)
31,000 to 51,999	37,068 (26.1)	2,964 (26.3)	60,313 (22.0)	4,769 (16.8)
18,000 to 30,999	23,965 (16.9)	2,439 (21.6)	66,591 (24.3)	6,510 (23.0)
Less than 18,000	11,374 (8.0)	1,875 (16.6)	60,392 (22.0)	9,831 (34.7)
Area-based deprivation (%)				
1 (Least deprived)	32,068 (22.6)	2,395 (21.2)	54,652 (19.9)	4,475 (15.8)
2	28,987 (20.4)	2,147 (19.0)	56,077 (20.4)	5,013 (17.7)
3	27,699 (19.5)	2,143 (19.0)	56,669 (20.6)	5,358 (18.9)
4	28,796 (20.3)	2,328 (20.6)	54,198 (19.7)	5,824 (20.5)
5 (Most deprived)	24,140 (17.0)	2,243 (19.9)	52,665 (19.2)	7,630 (26.9)
Body mass index (%)				
Under- or normal weight	58,879 (41.5)	3,931 (34.8)	84,819 (30.9)	7,471 (26.4)
Overweight	58,006 (40.9)	4,417 (39.1)	119,113 (43.4)	11,164 (39.4)
Obese	18,305 (12.9)	1,905 (16.9)	50,182 (18.3)	6,108 (21.6)
Severely obese	4,508 (3.2)	652 (5.8)	13,945 (5.1)	2,260 (8.0)
Morbidly obese	1,563 (1.1)	309 (2.7)	5,191 (1.9)	1,150 (4.1)
Physical activity (%)				
High	17,054 (12.0)	989 (8.8)	22,247 (8.1)	1,723 (6.1)
Moderate	93,762 (66.1)	7,026 (62.3)	184,013 (67.0)	16,947 (59.8)
Low	28,973 (20.4)	3,058 (27.1)	56,783 (20.7)	8,154 (28.8)
Alcohol intake (%)				
Safe drinking	62,500 (44.1)	5,284 (46.8)	122,977 (44.8)	13,939 (49.2)
Risky drinking	63,805 (45.0)	4,262 (37.8)	108,594 (39.5)	8,534 (30.1)
Smoking status (%)				
Never	87,651 (61.8)	6,230 (55.2)	147,093 (53.6)	13,374 (47.2)
Previous	43,969 (31.0)	3,799 (33.7)	95,961 (34.9)	9,896 (34.9)
Current	9,971 (7.0)	1,239 (11.0)	30,385 (11.1)	4,946 (17.5)
Fruit and vegetable intake per day (%)				
At least five a day	45,824 (32.3)	3,754 (33.3)	80,041 (29.1)	8,384 (29.6)
Less than five a day	95,953 (67.6)	7,525 (66.7)	194,149 (70.7)	19,902 (70.2)
Oily fish intake (%)				
At least once a week	83,289 (58.7)	6,304 (55.9)	148,183 (54.0)	14,057 (49.6)
Less than once per week	47,106 (33.2)	3,826 (33.9)	91,637 (33.4)	9,611 (33.9)
Never	11,125 (7.8)	1,119 (9.9)	32,863 (12.0)	4,430 (15.6)

*Table A 24 continued: Baseline characteristics for participants with depression and/ or low educational attainment (page 2 of 2)**

	No depression, high education (n=141,857, 31.1%)	**Depression, high education** (n=11,285, 2.5%)	**No depression, low education** (n=274,586, 60.2%)	**Depression, low education** (n=28,343, 6.2%)
Hypertension (%)	66,532 (46.9)	5,469 (48.5)	159,245 (58.0)	16,425 (58.0)
Diabetes (%)	4,490 (3.2)	535 (4.7)	12,607 (4.6)	1,875 (6.6)
High cholesterol levels (%)	16,047 (11.3)	1,705 (15.1)	43,651 (15.9)	5,854 (20.7)
Family history of cardiovascular disease (%)				
No	34,575 (24.4)	2,443 (21.6)	64,918 (23.6)	5,815 (20.5)
Yes	100,050 (70.5)	8,231 (72.9)	186,491 (67.9)	20,079 (70.8)
Family history of depression (%)				
No	116,684 (82.3)	7,752 (68.7)	214,372 (78.1)	18,975 (66.9)
Yes	12,868 (9.1)	2,420 (21.4)	20,368 (7.4)	4,863 (17.2)

* This table shows the baseline characteristics of participants without missing values in the education variable prior to imputation (n = 456,071). For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified.

IQR: Interquartile range

*Table A 25: Baseline characteristics for participants with depression and/ or high area-based deprivation (page 1 of 2)**

	No depression, less deprived (n=307,806, 66.2%)	**Depression, less deprived** (n=26,318, 5.7%)	**No depression, more deprived** (n=116,553, 25.1%)	**Depression, more deprived** (n=13,963, 3.0%)
Male (%)	139,139 (45.2)	7,579 (28.8)	53,735 (46.1)	4,658 (33.4)
Age (median [IQR])	58.0 [50.0, 63.0]	58.0 [51.0, 63.0]	56.0 [48.0, 62.0]	55.0 [48.0, 61.0]
Ethnicity (%)				
White	297,090 (96.5)	25,780 (98.0)	101,596 (87.2)	12,916 (92.5)
Other ethnic groups	9,500 (3.1)	455 (1.7)	13,850 (11.9)	933 (6.7)
Income (%)				
Greater than 100,000	16,381 (5.3)	732 (2.8)	4,899 (4.2)	216 (1.5)
52,000 to 100,000	64,327 (20.9)	3,747 (14.2)	14,251 (12.2)	813 (5.8)
31,000 to 51,999	75,364 (24.5)	5,938 (22.6)	22,111 (19.0)	1,796 (12.9)
18,000 to 30,999	65,521 (21.3)	6,299 (23.9)	25,363 (21.8)	2,667 (19.1)
Less than 18,000	41,278 (13.4)	5,758 (21.9)	31,155 (26.7)	6,008 (43.0)
Highest educational attainment (%)				
College or university degree	104,993 (34.1)	7,983 (30.3)	36,697 (31.5)	3,273 (23.4)
A levels, O levels, NVQ, or equivalent	156,772 (50.9)	13,710 (52.1)	51,817 (44.5)	6,457 (46.2)
None of the above	40,959 (13.3)	4,247 (16.1)	24,713 (21.2)	3,886 (27.8)
Body mass index (%)				
Under- or normal weight	108,525 (35.3)	7,853 (29.8)	37,510 (32.2)	3,702 (26.5)
Overweight	133,644 (43.4)	10,758 (40.9)	46,669 (40.0)	5,071 (36.3)
Obese	48,507 (15.8)	5,101 (19.4)	21,374 (18.3)	3,064 (21.9)
Severely obese	11,932 (3.9)	1,734 (6.6)	6,924 (5.9)	1,244 (8.9)
Morbidly obese	3,902 (1.3)	736 (2.8)	2,997 (2.6)	743 (5.3)
Physical activity (%)				
High	29,270 (9.5)	1,884 (7.2)	10,272 (8.8)	844 (6.0)
Moderate	203,922 (66.3)	16,056 (61.0)	76,332 (65.5)	8,093 (58.0)
Low	63,988 (20.8)	7,309 (27.8)	23,616 (20.3)	4,112 (29.4)
Alcohol intake (%)				
Safe drinking	135,856 (44.1)	12,769 (48.5)	52,851 (45.3)	6,723 (48.1)
Risky drinking	130,241 (42.3)	8,819 (33.5)	44,517 (38.2)	4,126 (29.5)
Smoking status (%)				
Never	180,621 (58.7)	14,268 (54.2)	57,971 (49.7)	5,629 (40.3)
Previous	102,833 (33.4)	9,138 (34.7)	39,269 (33.7)	4,729 (33.9)
Current	23,040 (7.5)	2,808 (10.7)	18,246 (15.7)	3,493 (25.0)
Fruit and vegetable intake per day (%)				
At least five a day	93,095 (30.2)	8,234 (31.3)	35,055 (30.1)	4,093 (29.3)
Less than five a day	214,175 (69.6)	18,045 (68.6)	80,639 (69.2)	9,772 (70.0)
Oily fish intake (%)				
At least once a week	174,332 (56.6)	14,118 (53.6)	61,180 (52.5)	6,561 (47.0)
Less than once per week	103,031 (33.5)	8,992 (34.2)	37,671 (32.3)	4,615 (33.1)
Never	28,928 (9.4)	3,079 (11.7)	15,935 (13.7)	2,567 (18.4)

*Table A 25 continued: Baseline characteristics for participants with depression and/ or high area-based deprivation (page 2 of 2)**

	No depression, less deprived (n=307,806, 66.2%)	**Depression, less deprived** (n=26,318, 5.7%)	**No depression, more deprived** (n=116,553, 25.1%)	**Depression, more deprived** (n=13,963, 3.0%)
Hypertension (%)	167,371 (54.4)	14,446 (54.9)	62,678 (53.8)	7,863 (56.3)
Diabetes (%)	10,921 (3.5)	1,361 (5.2)	6,660 (5.7)	1,121 (8.0)
High cholesterol levels (%)	42,342 (13.8)	4,703 (17.9)	18,638 (16.0)	3,026 (21.7)
Family history of cardiovascular disease (%)				
No	73,863 (24.0)	5,562 (21.1)	27,211 (23.3)	2,817 (20.2)
Yes	213,657 (69.4)	19,029 (72.3)	77,377 (66.4)	9,663 (69.2)
Family history of depression (%)				
No	249,142 (80.9)	18,338 (69.7)	86,885 (74.5)	8,725 (62.5)
Yes	24,242 (7.9)	4,800 (18.2)	9,418 (8.1)	2,574 (18.4)

* This table shows the baseline characteristics of participants without missing values in the deprivation variable prior to imputation (n = 464,640). For more information on the distribution of missing values see section 4.4.2 Investigation of missing data mechanism**.** Data are given as n (%) unless specified.

CSE: Certificate of secondary education, IQR: Interquartile range, NVQ: National vocational qualification

Table A 26: Hazard ratios (95% CI) of the association between depression and/ or low education and MCVE (imputed data)

	Unadjusted HR (95% CI)	**Adjusted HR (95% CI) Model 1**	**Adjusted HR (95% CI) Model 2**
No depression, high education	ref.	ref.	ref.
Depression alone	1.19 (1.01 – 1.40)	1.29 (1.09 – 1.51)	1.16 (0.99 – 1.37)
Low education alone	1.53 (1.45 – 1.61)	1.23 (1.16 – 1.31)	1.11 (1.05 – 1.18)
Depression and low education	1.85 (1.69 – 2.03)	1.69 (1.53 – 1.86)	1.35 (1.23 – 1.49)
No depression, low deprivation	ref.	ref.	ref.
Depression alone	1.07 (0.97 – 1.18)	1.21 (1.10 – 1.33)	1.10 (1.00 – 1.22)
High deprivation alone	1.19 (1.13 – 1.25)	1.18 (1.12 – 1.25)	1.07 (1.02 – 1.13)
Depression and high deprivation	1.73 (1.56 – 1.92)	1.87 (1.68 – 2.08)	1.46 (1.31 – 1.63)

Model 1: Age, sex, ethnicity, income, area-based deprivation (binary)

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval, HR: Hazard ratio, ref.: reference

Table A 27: Hazard ratios (95% CI) of the association between different measures of depression and MCVE in unadjusted, partially adjusted and fully adjusted models (complete cases)

	Unadjusted HR (95% CI)	**Adjusted HR (95% CI) Model 1**	**Adjusted HR (95% CI) Model 2**
Depression	1.13 (1.02 – 1.26)	1.26 (1.13 – 1.40)	1.14 (1.02 – 1.27)
Antidepressant use	1.15 (1.02 – 1.29)	1.28 (1.13 – 1.44)	1.14 (1.00 – 1.28)
Hospital diagnosis with depression	1.82 (1.35 – 2.46)	1.85 (1.37 – 2.50)	1.60 (1.18 – 2.16)
Self-reported depression	1.07 (0.94 – 1.23)	1.24 (1.09 – 1.42)	1.14 (0.99 – 1.31)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval, HR: Hazard ratio

Table A 28: Hazard ratios (95% CI) of the association between depression and stroke and MI (complete cases)

	Unadjusted HR (95% CI)	Adjusted HR (95% CI) Model 1	Adjusted HR (95% CI) Model 2
Stroke	1.24 (1.06 – 1.46)	1.30 (1.10 – 1.53)	1.21 (1.02 – 1.43)
Myocardial infarction	1.06 (0.93 – 1.22)	1.24 (1.08 – 1.43)	1.09 (0.95 – 1.26)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval, HR: Hazard ratio

Table A 29: Hazard ratios (95% CI) of the association between depression and MCVE, stroke, and MI, separately for men and women (complete cases)

	Major cardiovascular events	Stroke	Myocardial infarction
Women	1.28 (1.09 – 1.50)	1.17 (0.93 – 1.48)	1.37 (1.10 – 1.70)
Men	1.03 (0.88 – 1.20)	1.25 (0.98 – 1.59)	0.93 (0.77 – 1.12)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval

Table A 30: Hazard ratios (95% CI) of the association between depression and/ or male gender and MCVE (complete cases)

	Unadjusted HR (95% CI)	Adjusted HR (95% CI) Model 1	Adjusted HR (95% CI) Model 2
No depression, female	ref.	ref.	ref.
Depression alone	1.49 (1.28 – 1.74)	1.43 (1.22 – 1.67)	1.29 (1.10 – 1.51)
Male alone	2.65 (2.47 – 2.84)	2.64 (2.46 – 2.83)	2.39 (2.22 – 2.58)
Depression and male	3.16 (2.70 – 3.69)	2.99 (2.56 – 3.49)	2.44 (2.08 – 2.86)

CI: confidence interval, HR: Hazard ratio, ref.: reference

Table A 31: Hazard ratios (95% CI) of the association between depression and death from causes other than stroke or MI in unadjusted, partially adjusted and fully adjusted models (complete cases)

	Unadjusted HR (95% CI)	Adjusted HR (95% CI) Model 1	Adjusted HR (95% CI) Model 2
Depression	1.40 (1.28 – 1.52)	1.41 (1.29 – 1.54)	1.30 (1.19 – 1.43)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval, HR: Hazard ratio

Table A 32: Hazard ratios (95% CI) of the association between depression and/ or different comorbidities and MCVE (complete cases)

	Unadjusted HR (95% CI)	**Adjusted HR (95% CI) Model 1**	**Adjusted HR (95% CI) Model 2**
No depression, no hypertension	ref.	ref.	ref.
Depression alone	0.83 (0.65 – 1.07)	0.90 (0.70 – 1.15)	0.82 (0.64 – 1.05)
Hypertension alone	2.83 (2.62 – 3.04)	1.87 (1.73 – 2.02)	1.79 (1.66 – 1.94)
Depression and hypertension	3.44 (3.02 – 3.92)	2.55 (2.24 – 2.91)	2.22 (1.94 – 2.54)
No diabetes, no depression	ref.	ref.	ref.
Depression alone	1.11 (0.99 – 1.24)	1.25 (1.11 – 1.40)	1.14 (1.02 – 1.28)
Diabetes alone	2.38 (2.11 – 2.68)	1.58 (1.40 – 1.78)	1.33 (1.17 – 1.51)
Depression and diabetes	2.53 (1.85 – 3.46)	1.88 (1.37 – 2.57)	1.45 (1.06 – 2.00)
No depression, low cholesterol levels	ref.	ref.	ref.
Depression alone	1.02 (0.90 – 1.16)	1.16 (1.02 – 1.32)	1.07 (0.93 – 1.22)
High cholesterol levels alone	1.95 (1.81 – 2.10)	1.21 (1.12 – 1.31)	0.99 (0.91 – 1.07)
Depression and high cholesterol levels	2.47 (2.06 – 2.95)	1.75 (1.46 – 2.10)	1.29 (1.08 – 1.56)

Model 1: Age, sex, ethnicity, education, income, area-based deprivation

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval, HR: Hazard ratio, ref.: reference

Table A 33: Hazard ratios (95% CI) of the association between depression and/ or low socioeconomic status and MCVE (complete cases)

	Unadjusted HR (95% CI)	Adjusted HR (95% CI) Model 1	Adjusted HR (95% CI) Model 2
No depression, high education	ref.	ref.	ref.
Depression alone	1.19 (0.97 – 1.45)	1.28 (1.05 – 1.57)	1.16 (0.95 – 1.42)
Low education alone	1.46 (1.37 – 1.57)	1.21 (1.12 – 1.30)	1.10 (1.02 – 1.18)
Depression and low education	1.58 (1.38 – 1.80)	1.51 (1.32 – 1.73)	1.24 (1.07 – 1.42)
No depression, low deprivation	ref.	ref.	ref.
Depression alone	0.98 (0.85 – 1.12)	1.11 (0.97 – 1.28)	1.02 (0.89 – 1.17)
High deprivation alone	1.06 (0.99 – 1.14)	1.10 (1.02 – 1.19)	1.02 (0.94 – 1.10)
Depression and high deprivation	1.54 (1.30 – 1.81)	1.74 (1.47 – 2.05)	1.39 (1.18 – 1.65)

Model 1: Age, sex, ethnicity, income, area-based deprivation (binary)

Model 2: Model 1 + BMI, physical activity, alcohol intake, smoking, fruit and vegetable intake, oily fish intake, hypertension, cholesterol, diabetes, family history of CVD, and family history of depression

CI: Confidence interval, HR: Hazard ratio, ref.: reference

Table A 34: Results of formal tests for additive and multiplicative interaction between depression and selected socioeconomic factors and depression and selected comorbidities (complete cases)*

	RERI	Attributable proportion	Synergy index	Multiplicative interaction
Depression and sex	-0.24 (-0.66 to 0.17)	-0.10 (-0.28 to 0.08)	0.86 (0.65 to 1.13)	p = 0.03
Depression and hypertension	0.62 (0.29 to 0.94)	0.28 (0.15 to 0.40)	2.01 (1.31 to 3.09)	p < 0.01
Depression and diabetes	-0.02 (-0.51 to 0.48)	-0.01 (-0.35 to 0.33)	0.97 (0.33 to 2.84)	p = 0.81
Depression and high cholesterol levels	0.24 (-0.03 to 0.52)	0.19 (0.002 to 0.37)	5.67 (0.21 to 150.57)	p = 0.07
Depression and educational attainment	-0.03 (-0.31 to 0.25)	-0.02 (-0.25 to 0.20)	0.89 (0.30 to 2.64)	p = 0.77
Depression and area-based deprivation	0.36 (0.08 to 0.63)	0.25 (0.09 to 0.42)	9.98 (0.14 to 720.35)	p < 0.01

* Data are presented as estimate (95% confidence interval)

CI: Confidence interval, RERI: Relative excess risk due to interaction

A.III. ALSWH

*Table A 35: Participant characteristics at study phase 1, separately for those included and excluded from the analysis (page 1 of 2)**

	Included (n = 11,804, 82.9%)	**Excluded** (n = 2,443, 17.1%)	**p**
Age (median [IQR])	20.7 [19.5, 22.0]	20.5 [19.4, 21.9]	0.001
Marital status (%)			<0.001
Married/ de-facto	2,601 (22.0)	592 (24.2)	
Separated/ divorced/ widowed	89 (0.8)	45 (1.8)	
Single	9,062 (76.8)	1,788 (73.2)	
Missing value	52 (0.4)	18 (0.7)	
Lives alone (%)			0.361
Yes	714 (6.0)	138 (5.6)	
No	10,930 (92.6)	2,264 (92.7)	
Missing value	160 (1.4)	41 (1.7)	
Area of residence (%)			0.100
Metropolitan	6,502 (55.1)	1,361 (55.7)	
Rural	4,840 (41.0)	964 (39.5)	
Remote	436 (3.7)	109 (4.5)	
Missing value	26 (0.2)	9 (0.4)	
University degree (%)			<0.001
No	10,319 (87.4)	2,271 (93.0)	
Yes	1,426 (12.1)	150 (6.1)	
Missing value	59 (0.5)	22 (0.9)	
Ability to manage on income (%)			<0.001
Difficult/ impossible	2,065 (17.5)	559 (22.9)	
Difficult some of the time	3,852 (32.6)	854 (35.0)	
Easy/ not too bad	5,851 (49.6)	1,014 (41.5)	
Missing value	36 (0.3)	16 (0.7)	
Alcohol intake (%)			<0.001
Low risk	6,149 (52.1)	1048 (42.9)	
Non/ rarely	4,910 (41.6)	1,199 (49.1)	
Risky/ high risk	625 (5.3)	157 (6.4)	
Missing value	120 (1.0)	39 (1.6)	
Smoking status (%)			<0.001
Never-smoker	6,061 (51.3)	1,062 (43.5)	
Ex-smoker	1,729 (14.6)	356 (14.6)	
Current smoker	3,520 (29.8)	901 (36.9)	
Missing value	494 (4.2)	124 (5.1)	
Body mass index (kg/m²) (mean (SD))	22.9 (4.3)	22.5 (4.6)	0.003
BMI (WHO groups) (%)			<0.001
Underweight	955 (8.1)	279 (11.4)	
Acceptable weight	7,163 (60.7)	1,272 (52.1)	
Overweight	1,640 (13.9)	294 (12.0)	
Obese	689 (5.8)	123 (5.0)	
Missing value	1,357 (11.5)	475 (19.4)	

*Table A 35 continued: Participant characteristics at study phase 1, separately for those included and excluded from the analysis (page 2 of 2)**

	Included (n = 11,804, 82.9%)	**Excluded** (n = 2,443, 17.1%)	**p**
Diabetes (%)			0.003
Yes	108 (0.9)	38 (1.6)	
No	11,657 (98.8)	2,391 (97.9)	
Missing value	39 (0.3)	14 (0.6)	
Heart disease (%)			0.178
Yes	42 (0.4)	13 (0.5)	
No	11,717 (99.3)	2,416 (98.9)	
Missing value	45 (0.4)	14 (0.6)	
Hypertension (%)			0.019
Yes	573 (4.9)	149 (6.1)	
No	11,175 (94.7)	2,278 (93.2)	
Missing value	56 (0.5)	16 (0.7)	

* Data are given as n (%) unless specified

IQR: interquartile range, SD: standard deviation

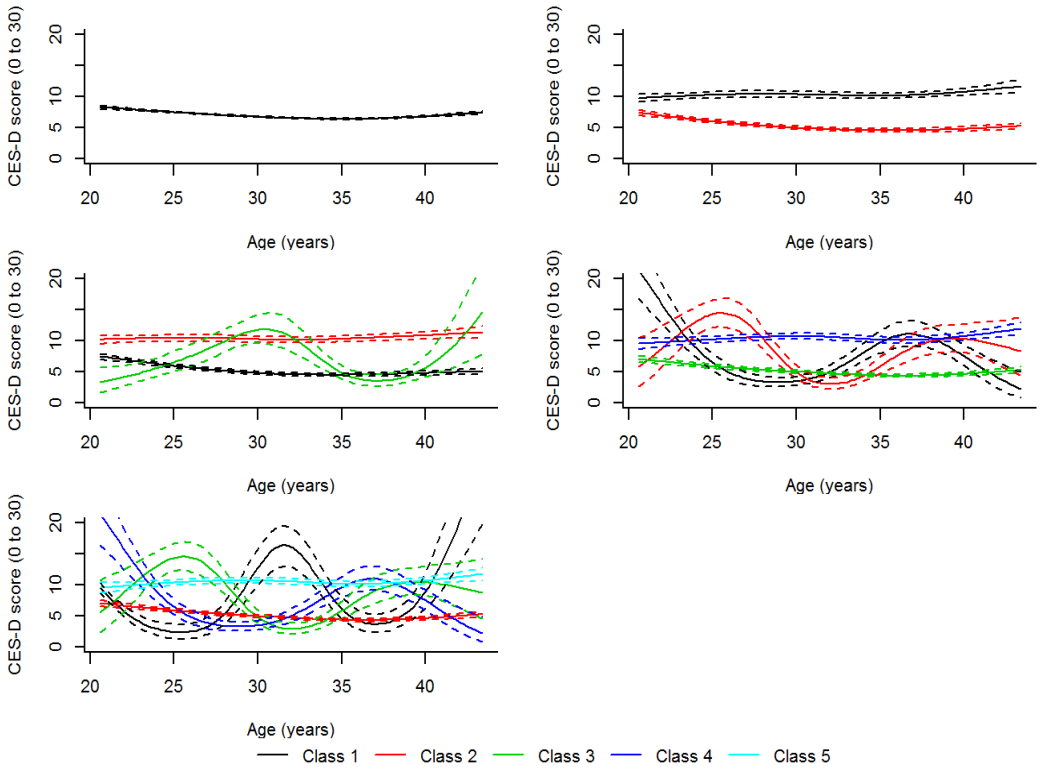


Figure A 1: Mean predicted trajectories of depressive symptoms in models with increasing numbers of groups

Table A 36: Numbers and proportion of women with hypertension, diabetes, and heart disease at any point during follow-up

	Stable low (n = 8,118)	Stable moderate (n= 3,558)	Fluctuating (n = 128)	p
Hypertension (%)	575 (7.1)	407 (11.4)	16 (12.5)	<0.01
Diabetes (%)	183 (2.3)	125 (3.5)	1 (0.8)	<0.01
Heart disease (%)	55 (0.7)	51 (1.4)	0 (0.0)	<0.01

* Data are given as n (%) unless specified

Table A 37: Model characteristics of models with increasing numbers of latent classes among those with at least two measures of depression

Number of groups	Number of iterations	Log-likelihood	Number of parameters	Discrete AIC	BIC	% class 1	% class 2	% class 3	% class 4	% class 5
1	17	-133885.3	14	266,249	267,900	100.0				
2	115	-133827.3	20	266,177	267,839	27.5	72.5			
3	164	-133779.7	26	266,104	267,800	1.0	27.5	71.5		
4	165	-133746.4	32	266,052	267,788	0.9	70.7	0.8	27.6	
5	331	-133718.2	38	266,050	267,788	46.0	49.8	1.1	0.5	2.7
6	500	Model did not reach convergence								
7	500	Model did not reach convergence								

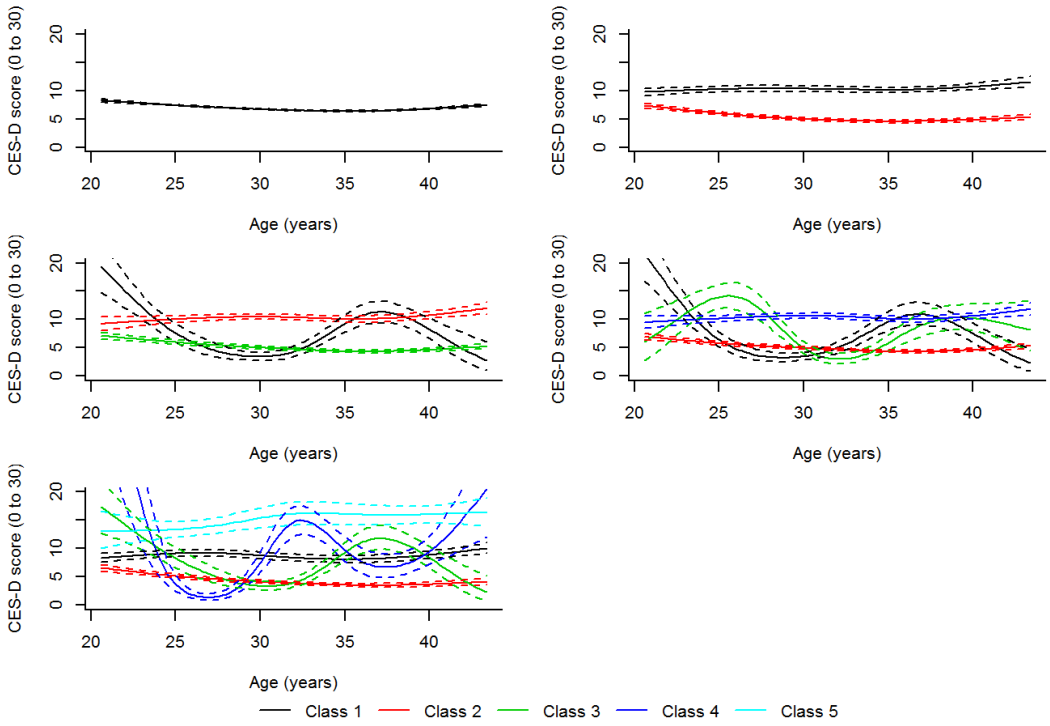


Figure A 2: Predicted trajectories of depressive symptoms in models with increasing numbers of latent classes in the analysis restricting the sample to women with at least two measures of depression

Table A 38: Model characteristics of models with increasing numbers of latent classes in the analysis using the unique item I felt depressed

Number of groups	Number of iterations	Log-likelihood	Number of parameters	Discrete AIC	BIC	% class 1	% class 2
1	25	-32,164.2	12	86,348	64,440	100.0	
2	402	-30,975.3	18	86,539	62,119	20.8	79.2
3	500	Model did not reach convergence					
4	500	Model did not reach convergence					

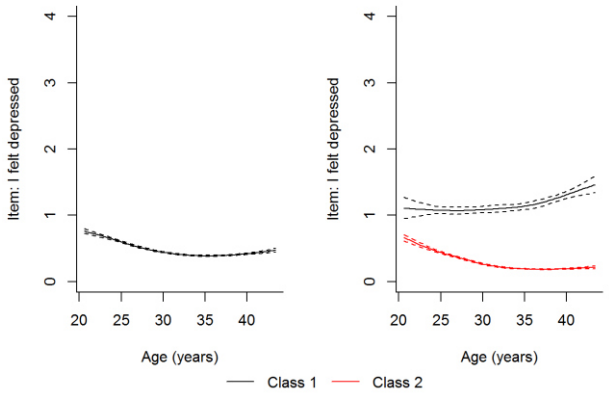


Figure A 3: Predicted trajectories of the unique item I felt depressed in models with increasing numbers of latent classes

Table A 39: Model characteristics of models with increasing numbers of latent classes in the analysis using the unique item I felt hopeful about the future

Number of groups	Number of iterations	Log-likelihood	Number of parameters	Discrete AIC	BIC	% class 1	% class 2
1	21	-58,764.0	12	122,265	117,640	100.0	
2	30	-58,466.1	18	122,417	117,100	52.9	47.1
3	500	Model did not reach convergence					
4	500	Model did not reach convergence					

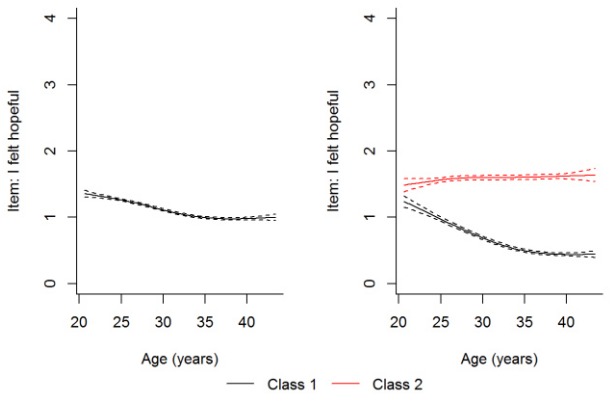


Figure A 4: Predicted trajectories of the unique item I felt hopeful about the future in models with increasing numbers of latent classes

A.IV. Whitehall II

*Table A 40: Baseline characteristics, separately for those included and excluded from the analyses on fatal coronary heart disease or non-fatal myocardial infarction during follow-up**

	Included (n = 10,187, 98.8%)	**Excluded** (n = 121, 1.2%)	p
Male (%)	6,803 (66.8)	92 (76.0)	0.040
Age (median [IQR])	44.24 [39.6, 50.3]	50.82 [46.5, 53.6]	<0.001
Ethnicity (%)			0.001
White	9,084 (89.2)	97 (80.2)	
Non-white	1,011 (9.9)	24 (19.8)	
Missing value	92 (0.9)	0 (0.0)	
Marital status (%)			0.506
Married/ cohabiting	7,523 (73.8)	85 (70.2)	
Other	2,627 (25.8)	35 (28.9)	
Missing value	37 (0.4)	1 (0.8)	
Occupational position (%)			0.001
Administrative	2,996 (29.4)	32 (26.4)	
Professional/ executive	4,898 (48.1)	45 (37.2)	
Clerical/ support	2,293 (22.5)	44 (36.4)	
Smoking status (%)			0.094
Never smoker	5,023 (49.3)	46 (38.0)	
Ex-smoker	3,232 (31.7)	49 (40.5)	
Current smoker	1,861 (18.3)	25 (20.7)	
Missing value	71 (0.7)	1 (0.8)	
Alcohol intake (%)			0.544
Risky drinking	2,394 (23.5)	25 (20.7)	
Safe drinking	7,701 (75.6)	94 (77.7)	
Missing value	92 (0.9)	2 (1.7)	
Physical activity (median [IQR])	3.0 [1.0, 5.0]	2.0 [1.0, 4.0]	0.275
Fruit and vegetable intake (%)			<0.001
Less than once a day	4,240 (41.6)	57 (47.1)	
Daily	4,231 (41.5)	46 (38.0)	
More than daily	1,686 (16.6)	15 (12.4)	
Missing value	30 (0.3)	3 (2.5)	
Body mass index (kg/ m2) (median [IQR])	24.2 [22.3, 26.3]	26.1 [23.8, 28.7]	<0.001
Diabetes (self-report or medication) (%)	95 (0.9)	5 (4.1)	0.002
Systolic blood pressure (median [IQR])	122.0 [113.0, 132.0]	124.0 [112.0, 134.0]	0.187
Diastolic blood pressure (median [IQR])	76.0 [70.0, 83.0]	79.0 [72.0, 88.0]	0.012
Total cholesterol levels (median [IQR])	5.9 [5.2, 6.7]	6.3 [5.6, 7.2]	<0.001
Psychological distress (median [IQR])	1.0 [0.0, 5.0]	1.0 [0.0, 4.0]	0.618
Psychological distress (binary) (%)			<0.001
No	7,372 (72.4)	73 (60.3)	
Yes	2,721 (26.7)	23 (19.0)	
Missing value	94 (0.9)	25 (20.7)	

* Data are given as n (%) unless specified; IQR: Interquartile range, SD: Standard deviation

*Table A 41: Baseline characteristics, separately for those included and excluded from the analyses on fatal or non-fatal stroke during follow-up**

	Included (n = 10,252, 99.5%)	**Excluded** (n = 56, 0.5%)	p
Male (%)	6,869 (67.0)	26 (46.4)	0.002
Age (median [IQR])	44.3 [39.6, 50.3]	48.9 [46.1, 51.6]	<0.001
Ethnicity (%)			<0.001
White	9,140 (89.2)	41 (73.2)	
Non-white	1,020 (9.9)	15 (26.8)	
Missing value	92 (0.9)	0 (0.0)	
Marital status (%)			0.351
Married/ cohabiting	7,571 (73.8)	37 (66.1)	
Other	2,643 (25.8)	19 (33.9)	
Missing value	38 (0.4)	0 (0.0)	
Occupational position (%)			<0.001
Administrative	3,022 (29.5)	6 (10.7)	
Professional/ executive	4,922 (48.0)	21 (37.5)	
Clerical/ support	2,308 (22.5)	29 (51.8)	
Smoking status (%)			0.451
Never smoker	5,044 (49.2)	25 (44.6)	
Ex-smoker	3,258 (31.8)	23 (41.1)	
Current smoker	1,878 (18.3)	8 (14.3)	
Missing value	72 (0.7)	0 (0.0)	
Alcohol intake (%)			0.128
Risky drinking	2,412 (23.5)	7 (12.5)	
Safe drinking	7,747 (75.6)	48 (85.7)	
Missing value	93 (0.9)	1 (1.8)	
Physical activity (median [IQR])	3.0 [1.0, 5.0]	2.0 [0.5, 2.5]	<0.001
Fruit and vegetable intake (%)			<0.001
Less than once a day	4,267 (41.6)	30 (53.6)	
Daily	4,255 (41.5)	22 (39.3)	
More than daily	1,700 (16.6)	1 (1.8)	
Missing value	30 (0.3)	3 (5.4)	
Body mass index (kg/ m2) (median [IQR])	24.2 [22.3, 26.7]	26.5 [23.2, 28.8]	<0.001
Diabetes (self-report or medication) (%)	98 (1.0)	2 (3.6)	0.191
Systolic blood pressure (median [IQR])	122.0 [113.0, 132.0]	126.5 [115.3, 140.3]	0.061
Diastolic blood pressure (median [IQR])	76.0 [70.0, 83.0]	81.0 [72.0, 90.0]	0.003
Total cholesterol levels (median [IQR])	5.9 [5.2, 6.7]	5.95 [5.4, 7.1]	0.187
Psychological distress (median [IQR])	1.0 [0.0, 5.0]	2.00 [0.0, 3.0]	0.590
Psychological distress (binary) (%)			<0.001
No	7,419 (72.4)	26 (46.4)	
Yes	2,737 (26.7)	7 (12.5)	
Missing value	96 (0.9)	23 (41.1)	

* Data are given as n(%) unless specified; IQR: Interquartile range, SD: Standard deviation

*Table A 42: Baseline characteristics, separately for those with and without fatal or non-fatal stroke during follow-up**

	No stroke (n = 10,059, 98.1%)	**Stroke** (n = 193, 1.9%)	p
Male (%)	6,731 (66.9)	138 (71.5)	0.206
Age (median [IQR])	44.2 [39.6, 50.2]	49.9 [43.5, 53.1]	<0.001
Ethnicity (%)			0.015
White	8,979 (89.3)	161 (83.4)	
Non-white	989 (9.8)	31 (16.1)	
Missing value	91 (0.9)	1 (0.5)	
Marital status (%)			0.299
Married/cohabiting	7,428 (73.8)	143 (74.1)	
Other	2,595 (25.8)	48 (24.9)	
Missing value	36 (0.4)	2 (1.0)	
Occupational position (%)			0.234
Administrative	2,971 (29.5)	51 (26.4)	
Professional/ executive	4,833 (48.0)	89 (46.1)	
Clerical/ support	2,255 (22.4)	53 (27.5)	
Smoking status (%)			0.020
Never smoker	4,956 (49.3)	88 (45.6)	
Ex-smoker	3,204 (31.9)	54 (28.0)	
Current smoker	1,827 (18.2)	51 (26.4)	
Missing value	72 (0.7)	0 (0.0)	
Alcohol intake (%)			0.402
Risky drinking	2,367 (23.5)	45 (23.3)	
Safe drinking	7,599 (75.5)	148 (76.7)	
Missing value	93 (0.9)	0 (0.0)	
Physical activity (median [IQR])	3.0 [1.0, 5.0]	2.0 [1.0, 4.0]	0.124
Fruit and vegetable intake (%)			0.247
Less than once a day	4,199 (41.7)	68 (35.2)	
Daily	4,164 (41.4)	91 (47.2)	
More than daily	1,666 (16.6)	34 (17.6)	
Missing value	30 (0.3)	0 (0.0)	
Body mass index (kg/ m2) (median [IQR])	24.2 [22.3, 26.3]	24.9 [22.7, 27.5]	0.006
Diabetes (self-report or medication) (%)	94 (0.9)	4 (2.1)	0.216
Systolic blood pressure (median [IQR])	122.0 [113.0, 132.0]	126.0 [117.0, 138.0]	<0.001
Diastolic blood pressure (median [IQR])	76.0 [70.0, 83.0]	80.0 [72.0, 88.0]	<0.001
Total cholesterol levels (median [IQR])	5.9 [5.2, 6.7]	5.9 [5.3, 6.7]	0.371
Psychological distress (median [IQR])	1.0 [0.0, 5.0]	2.0 [0.0, 6.3]	0.272
Psychological distress (binary) (%)			0.489
No	7,285 (72.4)	134 (69.4)	
Yes	2,679 (26.6)	58 (30.1)	
Missing value	95 (0.9)	1 (0.5)	
Person years at risk (mean (SD))	19.2 (6.4)	14.6 (5.5)	<0.001
Age at diagnosis/ end of follow-up (median [IQR])	64.0 [59.6, 70.3]	63.6 [57.2, 69.2]	0.021

* Data are given as n(%) unless specified; IQR: Interquartile range, SD: Standard deviation

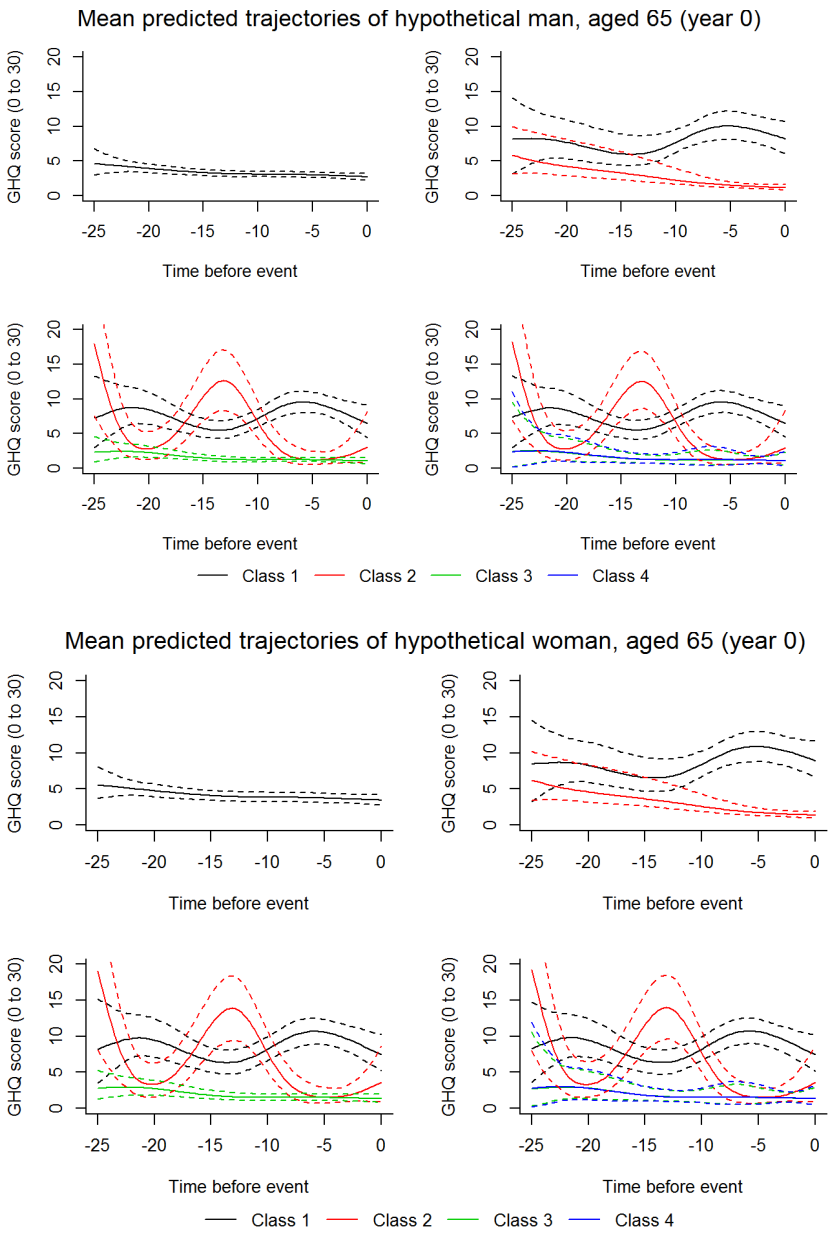


Figure A 5: Mean predicted trajectories of hypothetical men and women, aged 65 years at year 0 in models with increasing numbers of latent classes

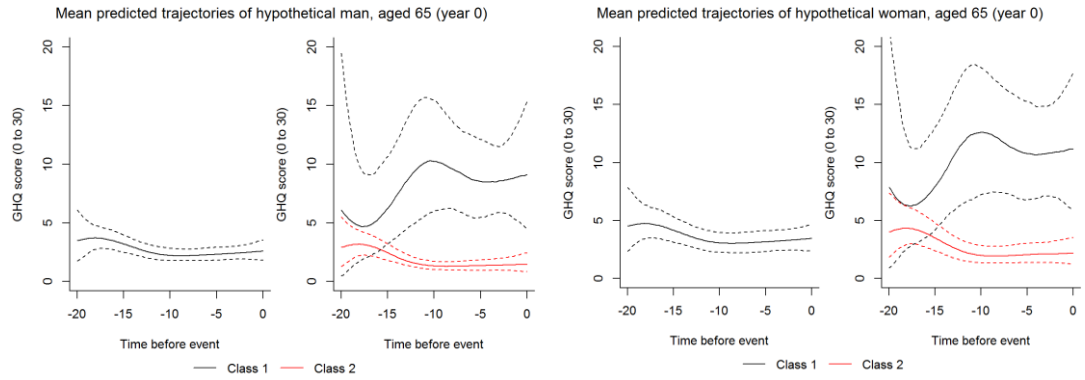


Figure A 6: Mean predicted trajectories of latent classes among those with fatal or non-fatal stroke